

Qy	366	GGCACCCAGCAGCGGATCGGCTGCCCTTGCCTGCAGAGGGCCCTGGGGGGCCCCCTTGGGG	425
Db	61	GGCACCCAGCAGCGCATCGGCTGGCCCTTGCCTGCAGAGGGCTGGGGGGCCCCCTTGGGG	120
Qy	426	CTGCGGCTGCCCGGAGACGACGAAGACCCGAGGAGCCCGCGCGGAGGGGACGCTTT	485
Db	121	CTGCGGCTGCCCGGAGACGACGAAGACCCGAGGAGCCCGCGCGGAGGGGACGCTTT	180
Qy	486	GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACCTTGGAGATGACC	545
Db	181	GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACCTTGGAGATGACC	240
Qy	546	GTGGGCAAGCCCCCGCAGACGCTCAACATCTCTGGTGGATACAGGCAGCAGTAACTTTGCA	605
Db	241	GTGGGCAAGCCCCCGCAGACGCTCAACATCTCTGGTGGATACAGGCAGCAGTAACTTTGCA	300
Qy	606	GTGGGTGCTGCCGCCACCCCTTCTGTGATCGCTACTACCAGAGCAGCTGTCCAGCAC	665
Db	301	GTGGGTGCTGCCGCCACCCCTTCTGTGATCGCTACTACCAGAGCAGCTGTCCAGCAC	360
Qy	666	TACCGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG	725
Db	361	TACCGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG	420
Qy	726	CTGGGCACCGACCTGGTAAAGCATCCCCATGGCCCCCAACGTCACTGTGCTGCCAACATTT	785
Db	421	CTGGGCACCGACCTGGTAAAGCATCCCCATGGCCCCCAACGTCACTGTGCTGCCAACATTT	480
Qy	786	GCTGCCATCACTGAATCAGACAAAGTTCTTCATCAAGGCTCCAACTGGGAAGGCATCCTG	845
Db	481	GCTGCCATCACTGAATCAGACAAAGTTCTTCATCAAGGCTCCAACTGGGAAGGCATCCTG	540
Qy	846	GGGCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT	905
Db	541	GGGCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT	600
Qy	906	CTGTGTAAGCAGACCCAGCTTCCCAACCTTTCTCCCTGCAGCTTTGTGGTGCCTGCTTC	965
Db	601	CTGTGTAAGCAGACCCAGCTTCCCAACCTTTCTCCCTGCAGCTTTGTGGTGCCTGCTTC	660
Qy	966	CCCTCAACCACTGCTGAAGTGTGGCCCTCTGTGCGAGGGAGCATGATCATTTGGAGCTATC	1025
Db	661	CCCTCAACCACTGCTGAAGTGTGGCCCTCTGTGCGAGGGAGCATGATCATTTGGAGCTATC	720
Qy	1026	GACCACTCGCTGTACACAGGCACTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTTAT	1085
Db	721	GACCACTCGCTGTACACAGGCACTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTTAT	780
Qy	1086	GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG	1145
Db	781	GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG	840
Qy	1146	TACAACATGACAAGACGATTTGTGACAGTGGCACCAACCTTCGTTTGGCCAAAGAAA	1205
Db	841	TACAACATGACAAGACGATTTGTGACAGTGGCACCAACCTTCGTTTGGCCAAAGAAA	900
Qy	1206	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCAGCTCCTCCACGGAGAAGTTCCCTGAT	1265
Db	901	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCAGCTCCTCCACGGAGAAGTTCCCTGAT	960
Qy	1266	GTTTCTTGCTAGGAGAGCAGTGTGTGCTGGCAAGCAGGCACCAACCTTGGAAACATTT	1325
Db	961	GTTTCTTGCTAGGAGAGCAGTGTGTGCTGGCAAGCAGGCACCAACCTTGGAAACATTT	1020
Qy	1326	TTCCCACTATCTCACTTACCTATNGGGTGAAGTTACCAACAGTCTTCCGCATCACC	1385
Db	1021	TTCCCACTATCTCACTTACCTATNGGGTGAAGTTACCAACAGTCTTCCGCATCACC	1080
Qy	1386	ATCCTCTCCGACGAATACCTCGCGCCAGTGGGAAGATGTGGCCACGCTCCCAAGACGACTGT	1445
Db	1081	ATCCTCTCCGACGAATACCTCGCGCCAGTGGGAAGATGTGGCCACGCTCCCAAGACGACTGT	1140
Qy	1446	TACAAGTTTGGCATCTCAGCTATCCACGGGCATGTTATGGGAGCTGTTATATCATGGAG	1505

Db	1141	TACAAAGTTTGGCATCTCAAGTATCCACGGCACTGTTATYGGGAGCTGTTATCATCGAG	1200
Qy	1506	GGCTTCTACGTTGTCTTTGATCGGGCCCGAAACGAATTGGCTTGTCTCGACGGCTTGC	1565
Db	1201	GGCTTCTACGTTGTCTTTGATCGGGCCCGAAACGAATTGGCTTGTCTCGACGGCTTGC	1260
Qy	1566	CATGTGCACGATGAGTTTCAGGACGGCAGGGTGGAGGCCCTTTTGTCTACCTTTGGACATG	1625
Db	1261	CATGTGCACGATGAGTTTCAGGACGGCAGGGTGGAGGCCCTTTTGTCTACCTTTGGACATG	1320
Qy	1626	GAAGACTCTGGCTACAAAGATPCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT	1685
Db	1321	GAAGACTCTGGCTACAAAGATPCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT	1380
Qy	1686	GTCATGGCTGGCATCTGGCGCCCTTTCATGCTGCCACTCTGGCTCATGTGTGTCTACGTG	1745
Db	1381	GTCATGGCTGGCATCTGGCGCCCTTTCATGCTGCCACTCTGGCTCATGTGTGTCTACGTG	1440
Qy	1746	CGCTGCCCTCGTGGCTCGGGCAGCAGCATGATGACTTTGTCTGATGACATCTCCCTGCTG	1805
Db	1441	CGCTGCCCTCGTGGCTCGGGCAGCAGCATGATGACTTTGTCTGATGACATCTCCCTGCTG	1500
Qy	1806	AAGTCAGCAGGCCATGTGGCAAGATAGAGATTCCCTTGGACACACACTCCGTGGTTCAT	1865
Db	1501	AAGTCAGCAGGCCATGTGGCAAGATAGAGATTCCCTTGGACACACACTCCGTGGTTCAT	1560
Qy	1866	CTTTTGGTCCAAAGTAGGAGACACAGATGGCACCTCTGGCCAGAGCACCTCAGACCCCTCC	1925
Db	1561	CTTTTGGTCCAAAGTAGGAGACACAGATGGCACCTCTGGCCAGAGCACCTCAGACCCCTCC	1620
Qy	1926	CCACCCACCAATGCCTCTGCCTTCATGGAGAAGGAAAGGCTGGCAAGGTGGGTTCACAG	1985
Db	1621	CCACCCACCAATGCCTCTGCCTTCATGGAGAAGGAAAGGCTGGCAAGGTGGGTTCACAG	1680
Qy	1986	GGACTGTACCTGTAGGAACAGAAAGAGAAAGAACACTCTGCTGGCGGGAATACT	2045
Db	1681	GGACTGTACCTGTAGGAACAGAAAGAGAAAGAACACTCTGCTGGCGGGAATACT	1740
Qy	2046	CTTGGTCACTTCAATTTAAGTCGGGAAATCTGCTGCTTGAACCTTCAGCCCTGAACCT	2105
Db	1741	CTTGGTCACTTCAATTTAAGTCGGGAAATCTGCTGCTTGAACCTTCAGCCCTGAACCT	1800
Qy	2106	TTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATTCCTTTCTTTAGTTTTCAGAA	2165
Db	1801	TTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATTCCTTTCTTTAGTTTTCAGAA	1860
Qy	2166	GTACTGGCATCACACGACGGTTACCTTTGGCGTGTGTCCCTGTGGGTACCCTGGCAGAGAAG	2225
Db	1861	GTACTGGCATCACACGACGGTTACCTTTGGCGTGTGTCCCTGTGGGTACCCTGGCAGAGAAG	1920
Qy	2226	AGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCATATT	2285
Db	1921	AGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCATATT	1980
Qy	2286	TGCTTTTAGACACAGGACTGTATAAACAAAGCCTTAACATTTGGTGTCAAAGATTTGCCTCTTGA	2345
Db	1981	TGCTTTTAGACACAGGACTGTATAAACAAAGCCTTAACATTTGGTGTCAAAGATTTGCCTCTTGA	2040
Qy	2346	ATT 2348	
Db	2041	ATT 2043	

RESULT 2
US-03-795-847-3
: Sequence 3, Application US/09759847
: Patent No. US20010018208A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Bienkowski, Michael J.
: APPLICANT: Heinrichson, Robert L.
: APPLICANT: Parodi, Luis A.

APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795,847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-847-3

Query Match 86.9%; Score 2039.8; DB 10; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	966	CCCCTCAACAGCTCTGAAGTGGCTGGCTCTGTCTGGAGGAGCATGATCATTTGGAGGTATC	1025
DB	661	CCCCCTCAACAGCTCTGAAGTGGCTGGCTCTGTCTGGAGGAGCATGATCATTTGGAGGTATC	720
QY	1026	GACCACTCGCTGTACACAGGCACTCTCTGGTATACACCCATCGGGGGAGTGGTATTAT	1085
DB	721	GACCACTCGCTGTACACAGGCACTCTCTGGTATACACCCATCGGGGGAGTGGTATTAT	780
QY	1086	GAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	1145
DB	781	GAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	840
QY	1146	TACAACATATGACAAGAGCATTTGTGGACAGTGGACACCAACCTTCTGTTTGGCCCAAGAA	1205
DB	841	TACAACATATGACAAGAGCATTTGTGGACAGTGGACACCAACCTTCTGTTTGGCCCAAGAA	900
QY	1206	GTGTTTGAAGCTGCAGTCAAATCCATCAAGGAGGCTCTCCACGGAGAGTTCCTCTGAT	1265
DB	901	GTGTTTGAAGCTGCAGTCAAATCCATCAAGGAGGCTCTCCACGGAGAGTTCCTCTGAT	960
QY	1266	GGTTTCTGGCTAGGAGAGCAGCTGGTGTCTGGCAAGCAGGACACCCCTTGGAAACATT	1325
DB	961	GGTTTCTGGCTAGGAGAGCAGCTGGTGTCTGGCAAGCAGGACACCCCTTGGAAACATT	1020
QY	1326	TTCCAGTCTCATCTACCTTACCTTAATGGTGAGGTTTACCAACAGTCTTCCGCATCACC	1385
DB	1021	TTCCAGTCTCATCTACCTTACCTTAATGGTGAGGTTTACCAACAGTCTTCCGCATCACC	1080
QY	1386	ATCTCTTCCGAGCAATACCTCGGGCCAGTGGAGATGTGGCCACGTCCTCAAGACGACTGT	1445
DB	1081	ATCTCTTCCGAGCAATACCTCGGGCCAGTGGAGATGTGGCCACGTCCTCAAGACGACTGT	1140
QY	1446	TACAAGTTTGGCATCTCACAGTCACTCCAGGCACTGTTATGGAGCTGTTATCATGGAG	1505
DB	1141	TACAAGTTTGGCATCTCACAGTCACTCCAGGCACTGTTATGGAGCTGTTATCATGGAG	1200
QY	1506	GGCTTCTACGTTGTCTTTGATCGGGCCGAAACAGAAATTTGGCTTTGCTGTCAGCGCTTC	1565
DB	1201	GGCTTCTACGTTGTCTTTGATCGGGCCGAAACAGAAATTTGGCTTTGCTGTCAGCGCTTC	1260
QY	1566	CATGTGCAGCATGATTCAGGAGGCGGAGGTTGGAGGCGCTTTTGTCTACCTTTGGACATG	1625
DB	1261	CATGTGCAGCATGATTCAGGAGGCGGAGGTTGGAGGCGCTTTTGTCTACCTTTGGACATG	1320
QY	1626	GAAGACTGTGGCTACAAATTCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT	1685
DB	1321	GAAGACTGTGGCTACAAATTCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT	1380
QY	1686	GTATGGCTGCCATCTGCGCCCTTCTTCATGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG	1745
DB	1381	GTATGGCTGCCATCTGCGCCCTTCTTCATGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG	1440
QY	1746	CGCTGCCCTCGCTGCTGGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG	1805
DB	1441	CGCTGCCCTCGCTGCTGGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG	1500
QY	1806	AAGTAGGAGGCGCCATGCGCAGAAGATAGAGATTCCTCCCTGGACACACCTCCCTGGTGTCA	1865
DB	1501	AAGTAGGAGGCGCCATGCGCAGAAGATAGAGATTCCTCCCTGGACACACCTCCCTGGTGTCA	1560
QY	1866	CTTTGGTCAACAAGTAGGAGACACAGATGGACCTGTGGCCAGACACCTCAGGACCTCC	1925
DB	1561	CTTTGGTCAACAAGTAGGAGACACAGATGGACCTGTGGCCAGACACCTCAGGACCTCC	1620
QY	1926	CCACCCACCAATGCTCTGCTTGTATGAGAGGAGAAAGGCTGGCAAGTGGGTGCTCCAG	1985
DB	1621	CCACCCACCAATGCTCTGCTTGTATGAGAGGAGAAAGGCTGGCAAGTGGGTGCTCCAG	1680
QY	1986	GGACTGTACCTGTAGGAAACAGAAAGAGAAAGCACTCTGCTGGCGGGAATACT	2045
DB	1681	GGACTGTACCTGTAGGAAACAGAAAGAGAAAGCACTCTGCTGGCGGGAATACT	1740
QY	2046	CTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTGCTTGAACCTTTCAGCGTGAACCT	2105

QY 1566 CATGTGCACGATGAGTTTCAGGACGGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1625
|||||
Db 1261 CATGTGCACGATGAGTTTCAGGACGGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1320
QY 1626 GAAGACTGTGGCTACAACTTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1685
|||||
Db 1321 GAAGACTGTGGCTACAACTTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1380
QY 1686 GTCATGGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGGCTCATGTGTGTCAAGTGG 1745
|||||
Db 1381 GTCATGGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGGCTCATGTGTGTCAAGTGG 1440
QY 1746 CGCTGCTCCGCTGCGTGGCCAGCAGCATGATGACTTTGTGTGATGACATCTCCCTGCTG 1805
|||||
Db 1441 CGCTGCTCCGCTGCGTGGCCAGCAGCATGATGACTTTGTGTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGAAGAGGCCATGGGCAGAAATACAGATTCCCTTGGACCACACCTCCGTGTTTCA 1865
|||||
Db 1501 AAGTGAAGAGGCCATGGGCAGAAATACAGATTCCCTTGGACCACACCTCCGTGTTTCA 1560
QY 1866 TTCTTGTCTACAAGTAGGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCTCC 1925
|||||
Db 1561 TTCTTGTCTACAAGTAGGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCTCC 1620
QY 1926 CCACCCACCAATGCTCTGCTTGTATGGAGAAAGGCTGGCAAGTGGGTTTCCAG 1985
|||||
Db 1621 CCACCCACCAATGCTCTGCTTGTATGGAGAAAGGCTGGCAAGTGGGTTTCCAG 1680
QY 1986 GGACTGTACTGTAGGAACAGAAAGAGAAAGAACGACTCTGCTGGCGGAATACT 2045
|||||
Db 1681 GGACTGTACTGTAGGAACAGAAAGAGAAAGAACGACTCTGCTGGCGGAATACT 1740
QY 2046 TTCTTGTCTACAATTTAAGTCGGAAATCTGCTGCTTGAACCTTCAGCCCTGAACCT 2105
|||||
Db 1741 TTCTTGTCTACAATTTAAGTCGGAAATCTGCTGCTTGAACCTTCAGCCCTGAACCT 1800
QY 2106 TTGTCCACCATCTCTTAAATTTCCAAACCAAGATTTCTTCTTTTCTTAGTTTCAGAA 2165
|||||
Db 1801 TTGTCCACCATCTCTTAAATTTCCAAACCAAGATTTCTTCTTTTCTTAGTTTCAGAA 1860
QY 2166 GTACTGGCATCACGCGAGTTACCTTGGCGTGTGTCCTTGGTACCTTGGCAGAGAAG 2225
|||||
Db 1861 GTACTGGCATCACGCGAGTTACCTTGGCGTGTGTCCTTGGTACCTTGGCAGAGAAG 1920
QY 2226 AGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGGATGCACAGTTTGTATT 2285
|||||
Db 1921 AGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGGATGCACAGTTTGTATT 1980
QY 2286 TGCCTTAGACAGGAGCTGTATAACAAAGCCTTAACATTTGGTGCAAGATTTGCTCTTGA 2345
|||||
Db 1981 TGCCTTAGACAGGAGCTGTATAACAAAGCCTTAACATTTGGTGCAAGATTTGCTCTTGA 2040
QY 2346 ATT 2348
Db 2041 ATT 2043

RESULT 4

US-09-794-748-3
; Sequence 3, Application US/09794748
; Patent No. US200307315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-794-748-3

Query Match 86.9%; Score 2039.8; DB 10; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 306 ATGGCCCAAGCCCTGCGCTGGCTCTCTGCTGTGATGGGCGGGAGTGTGCTGCCCCAC 365
Db 1 ATGGCCCAAGCCCTGCGCTGGCTCTCTGCTGTGATGGGCGGGAGTGTGCTGCCCCAC 60
QY 366 GGCACCCAGCAGCGCATCGGCTGCCCTGCGCAGCGGCTGGGGGCGGCCCTCGGG 425
Db 61 GGCACCCAGCAGCGCATCGGCTGCCCTGCGCAGCGGCTGGGGGCGGCCCTCGGG 120
QY 426 CTGCGCTGCGCCCGGAGACCGAGAGCCCGAGAGCCCGGCGGAGGCGGCGGCGGCGGCTTT 485
Db 121 CTGCGCTGCGCCCGGAGACCGAGAGCCCGAGAGCCCGGCGGAGGCGGCGGCGGCGGCTTT 180
QY 486 GTGGAGATGGTGACAACTGTAGGGGGAAGTCGGGGAGGGCTACTACGTGGAGATGACC 545
Db 181 GTGGAGATGGTGACAACTGTAGGGGGAAGTCGGGGGCAAGTCGGGGAGGCGGCTACTACGTGGAGATGACC 240
QY 546 GTGGGAGCCCCCGCAGAGCGCTCAACATCCTGTGTGATACAGCAGCAGCTAACTTTTGA 605
Db 241 GTGGGAGCCCCCGCAGAGCGCTCAACATCCTGTGTGATACAGCAGCAGCTAACTTTTGA 300
QY 606 GTGGGTGCTGCCCGCCACCCCTTCTCTGATCGCTACTACAGAGGCGAGCTGTCCAGCACA 665
Db 301 GTGGGTGCTGCCCGCCACCCCTTCTCTGATCGCTACTACAGAGGCGAGCTGTCCAGCACA 360
QY 666 TACCGGAGCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGAAGTGGGAAGGGAG 725
Db 361 TACCGGAGCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGAAGTGGGAAGGGAG 420
QY 726 CTGGGACCGACCTGGTAAAGCATCCCCCATGGCCCAACGCTCACTGTGCGTGCCAACTT 785
Db 421 CTGGGACCGACCTGGTAAAGCATCCCCCATGGCCCAACGCTCACTGTGCGTGCCAACTT 480
QY 786 GCTGCCATCACTGAATCAGACAAAGTTCTTTCATCAACGGCTCCCACTGGGAAGGCATCTG 845
Db 481 GCTGCCATCACTGAATCAGACAAAGTTCTTTCATCAACGGCTCCCACTGGGAAGGCATCTG 540
QY 846 GGGCTGGCTATGCTGAGATTGCCAGGCTGAGGCTCCCTGGAGCCTTTCTTTGACTCT 905
Db 541 GGGCTGGCTATGCTGAGATTGCCAGGCTGAGGCTCCCTGGAGCCTTTCTTTGACTCT 600
QY 906 CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCTGCGAGCTTTTGTGTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCTGCGACCTTTGTGTGTGGCTTC 660
QY 966 CCCTCAACAGTCTGAAAGTGTGGCTCTGTCTCGGAGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCTCAACAGTCTGAAAGTGTGGCTCTGTCTCGGAGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACTCGGTACACAGGAGCTCTCTGTGTATACACCCATCCGGGGGAGTGGTATTAT 1085
Db 721 GACCACTCGGTACACAGGAGCTCTCTGTGTATACACCCATCCGGGGGAGTGGTATTAT 780

QY 606 GTGGGTGCTGCCCCACACCCCTTCTCGCATCGCTACTACAGAGGAGCTGTCCAGCACA 665
DB 301 GTGGGTGCTGCCCCACACCCCTTCTCGCATCGCTACTACAGAGGAGCTGTCCAGCACA 360
QY 666 TACCGGAGCTCCGGGAAGGCTGTATGTATGTCCTTACACCCAGGCGCAAGTGGGAAGGGAG 725
DB 361 TACCGGAGCTCCGGGAAGGCTGTATGTATGTCCTTACACCCAGGCGCAAGTGGGAAGGGAG 420
QY 726 CTGGGACCGACCTGTAAAGCATCCCCCATGGCCCAAGCTCACTGTGGTGCACACATT 795
DB 421 CTGGGACCGACCTGTAAAGCATCCCCCATGGCCCAAGCTCACTGTGGTGCACACATT 480
QY 786 GTGGCATCACTGAATCAGACAAGTTCTTCATCAAGGGGTCCAACTGGGAAGGCATCCCTG 845
DB 481 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAAGGGGTCCAACTGGGAAGGCATCCCTG 540
QY 846 GGGCTGGCTATGCTGAGATGGCAGGCTGACGACTCCCTGGAGGCTTTCTTTGACTCT 905
DB 541 GGGCTGGCTATGCTGAGATGGCAGGCTGACGACTCCCTGGAGGCTTTCTTTGACTCT 600
QY 906 CTGGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGTGGCTTC 965
DB 601 CTGGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGTGGCTTC 660
QY 966 CCCTCAAGCAGTCTGAAGTGTGGCTCTGTTCGGAGGAGCATGATCATTTGGAGGTATC 1025
DB 661 CCCTCAAGCAGTCTGAAGTGTGGCTCTGTTCGGAGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACTCGCTGACACAGCAGTCTCTGGTATACACCCATCCGGGGAGTGGTATTAT 1085
DB 721 GACCACTCGCTGACACAGCAGTCTCTGGTATACACCCATCCGGGGAGTGGTATTAT 780
QY 1086 GAGGTATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAATGGACTGCAAGGAG 1145
DB 781 GAGGTATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAATGGACTGCAAGGAG 840
QY 1146 TACAACATATCACAAGAGCATTTGGAGACAGTGGCCACCAACCTCTGTTTGGCCCAAGAAA 1205
DB 841 TACAACATATCACAAGAGCATTTGGAGACAGTGGCCACCAACCTCTGTTTGGCCCAAGAAA 900
QY 1206 GTGTTTGAAGTGCATCAATCCATCAAGGACGCTCTCCACGAGAGAGTTCCCTGAT 1265
DB 901 GTGTTTGAAGTGCATCAATCCATCAAGGACGCTCTCCACGAGAGAGTTCCCTGAT 960
QY 1266 GGTTCCTGGCTAGCAGCAGCTGTGTGCTGCAAGCAGGACCAACCTTGGACATT 1325
DB 961 GGTTCCTGGCTAGCAGCAGCTGTGTGCTGCAAGCAGGACCAACCTTGGACATT 1020
QY 1326 TTCCCACTCATCTCACTTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCATCACC 1385
DB 1021 TTCCCACTCATCTCACTTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCATCACC 1080
QY 1386 ATCTTTCCGAGCAATACCTTCGGCCAGTGGGAAGATGTGGCCACGCTCCCAAGACGACTGT 1445
DB 1081 ATCTTTCCGAGCAATACCTTCGGCCAGTGGGAAGATGTGGCCACGCTCCCAAGACGACTGT 1140
QY 1446 TACAACCTTTGCCATCTCACAGTCAATCCAGGGCACTGTTATGGGAGCTGTTATCATGGAG 1505
DB 1141 TACAACCTTTGCCATCTCACAGTCAATCCAGGGCACTGTTATGGGAGCTGTTATCATGGAG 1200
QY 1506 GGTTCCTACGCTTCTTTGATCGGGCCGCAAAACGAATTGGCTTTGCTGTCAGCGCTTGC 1565
DB 1201 GGTTCCTACGCTTCTTTGATCGGGCCGCAAAACGAATTGGCTTTGCTGTCAGCGCTTGC 1260
QY 1566 CATGTGCAGATGAGTTTCAGGACGCGGCTGGGAAGGCCCTTTTGTCACTTTGGACATG 1625
DB 1261 CATGTGCAGATGAGTTTCAGGACGCGGCTGGGAAGGCCCTTTTGTCACTTTGGACATG 1320
QY 1626 GAGACTGTGGCTACAAACATTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1685
DB 1321 GAGACTGTGGCTACAAACATTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1380

QY 1686 GTCATGGCTGCCATCTGCGCCCTTTCATGCTGCCACTCTGCCTCATGTGTGTGTCAGTGG 1745
DB 1381 GTCATGGCTGCCATCTGCGCCCTTTCATGCTGCCACTCTGCCTCATGTGTGTGTCAGTGG 1440
QY 1746 CGCTGGCTCGGCTGCGCTGCGCCAGCAGCATGATGACTTTTCTGCTGATGACATCTCCCTGCTG 1805
DB 1441 CGCTGGCTCGGCTGCGCTGCGCCAGCAGCATGATGACTTTTCTGCTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGAAGGAGGCCCATTTGGGAGAGATAGAGATTTCCCTGCGACACACCTCCGTTGGTTCA 1865
DB 1501 AAGTGAAGGAGGCCCATTTGGGAGAGATAGAGATTTCCCTGCGACACACCTCCGTTGGTTCA 1560
QY 1866 CTTTGGTTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCAGCTTCAGACCCCTCC 1925
DB 1561 CTTTGGTTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCAGCTTCAGACCCCTCC 1620
QY 1926 CCACCCACCAATGCTCTGCTTGTATGGAGAAGGAAAGGCTGGCAAGGTGGTTCAG 1985
DB 1621 CCACCCACCAATGCTCTGCTTGTATGGAGAAGGAAAGGCTGGCAAGGTGGTTCAG 1680
QY 1986 GGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAGCACTCTGCTGCCGGAATACT 2045
DB 1681 GGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAGCACTCTGCTGCCGGAATACT 1740
QY 2046 CTTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTGTTGAAACTTCAGCCCTGAACCT 2105
DB 1741 CTTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTGTTGAAACTTCAGCCCTGAACCT 1800
QY 2106 TTGTCACCACTTCTTTAAATTTCCAAACCAAGATATTTCTTTTCTTAGTTTCAGAA 2165
DB 1801 TTGTCACCACTTCTTTAAATTTCCAAACCAAGATATTTCTTTTCTTAGTTTCAGAA 1860
QY 2166 GTACTGSCATCACACGAGTTACCTTGGCGTGTGTCCTGTGTACCCCTGGCAGAGAAG 2225
DB 1861 GTACTGSCATCACACGAGTTACCTTGGCGTGTGTCCTGTGTACCCCTGGCAGAGAAG 1920
QY 2226 AGACCAAGCTTGTTCCTGCTGCGCAAGTCAGTAGGAGAGGATGCACAGTTTGTCTATT 2285
DB 1921 AGACCAAGCTTGTTCCTGCTGCGCAAGTCAGTAGGAGAGGATGCACAGTTTGTCTATT 1980
QY 2286 TGCTTTAGACAGAGGAGCTGTATAACCAAGCTTAACATTTGGTGCAGAGATTGCTCTTGA 2345
DB 1981 TGCTTTAGACAGAGGAGCTGTATAACCAAGCTTAACATTTGGTGCAGAGATTGCTCTTGA 2040
QY 2346 ATT 2348
DB 2041 ATT 2043

RESULT 6
US-09-681-442-3
; Sequence 3, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Blenkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/681,442
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594

306	Qy	ATGCCCCAAGCCCTGCCCTGGCTCTCTGTGTGATGGCGGGGAGTGTCTGCTTCCGCAC	365
1	Db	ATGGCCCAAGCCCTGCCCTGGCTCTCTGTGTGATGGCGGGGAGTGTCTGCTTCCGCAC	60
366	Qy	GGCACCAGCAGCGATCCGGCTGCCCTTCGCGAGCGGCTTGGGGGGCGCCCTCGGGG	425
61	Db	GGCACCAGCAGCGATCCGGCTGCCCTTCGCGAGCGGCTTGGGGGGCGCCCTCGGGG	120
426	Qy	CTGCGGCTGCCCGGAGACCGACGAAGACCCGAGAGCGCCGCCCGGAGGGGACGCTTT	485
121	Db	CTGCGGCTGCCCGGAGACCGACGAAGACCCGAGAGCGCCGCCCGGAGGGGACGCTTT	180
486	Qy	GTGGAGATGGTGACAACTGAGGGCAAGTCGGGGCAGGCTACTAGTGTGGAGTACCC	545
181	Db	GTGGAGATGGTGGAACCTGAGGGCAAGTCGGGGCAGGCTACTAGTGTGGAGTACCC	240
546	Qy	GTGGCAGCCCCCGCAGACGCTCAACATCTCGTGTGATACAGCAGCAGTAACCTTTGCA	605
241	Db	GTGGCAGCCCCCGCAGACGCTCAACATCTCGTGTGATACAGCAGCAGTAACCTTTGCA	300
606	Qy	GTGGGTGCTGCCCGCCACCCCTTCCTTGATCTGCTACTACAGAGGCAGCTGTCCAGCACA	665
301	Db	GTGGGTGCTGCCCGCCACCCCTTCCTTGATCTGCTACTACAGAGGCAGCTGTCCAGCACA	360
666	Qy	TACCGGGACCTCCGGAAGGTTGTATGTGCCCTTACACCCAGGGCAAGTGGGAAGGGGAG	725
361	Db	TACCGGGACCTCCGGAAGGTTGTATGTGCCCTTACACCCAGGGCAAGTGGGAAGGGGAG	420
726	Qy	CTGGCAGCGACCTGTGTAGCATCCCCATGGCCCCAACGTCACTGTGCTGGCCAACTTT	785
421	Db	CTGGCAGCGACCTGTGTAGCATCCCCATGGCCCCAACGTCACTGTGCTGGCCAACTTT	480
786	Qy	GCTGCCATCACTGAATCAGACAAGTTCTTCACTCAACGGCTCCAACCTGGGAAGGATCCCTG	845
481	Db	GCTGGCATCACTGAATCAGACAAGTTCTTCACTCAACGGCTCCAACCTGGGAAGGATCCCTG	540
846	Qy	GGGTGGGCTATGCTGAGATTGCCAGGCTGAGCACTCCCTGGAGCCTTTCTTTTGACTCT	905
541	Db	GGGTGGGCTATGCTGAGATTGCCAGGCTGAGCACTCCCTGGAGCCTTTCTTTTGACTCT	600
906	Qy	CTGTGTAAGCAGACCCACGTTCCCAACCTTCTCTCCCTGCAGCTTTGTGGTGGCTGTC	965
601	Db	CTGTGTAAGCAGACCCACGTTCCCAACCTTCTCTCCCTGCAGCTTTGTGGTGGCTGTC	660
966	Qy	CCCTCAACCACTCTCAAGTGTGGCTCTGTGGGAGGAGCATGATCATTTGGAGGTATC	1025
661	Db	CCCTCAACCACTCTCAAGTGTGGCTCTGTGGGAGGAGCATGATCATTTGGAGGTATC	720
1026	Qy	GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCCGGGGGAGTGGTATTAT	1085
721	Db	GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCCGGGGGAGTGGTATTAT	780
1086	Qy	GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG	1145
781	Db	GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG	840
1146	Qy	TACAACTATGACAAGAGCATTTGGGACGTGGCACCCACCAACCTTCGTTGCCCAAGAAA	1205
841	Db	TACAACTATGACAAGAGCATTTGGGACGTGGCACCCACCAACCTTCGTTGCCCAAGAAA	900


```
|||||
Db 1981 TGCTTTAGACAGCAGGAGTGTATATAACAGCCTTAACATTGGTGCAAGATTGGCTCTCTTGA 2040
QY 2346 ATT 2348
|||
Db 2041 ATT 2043

RESULT 7
US-09-969-671A-1
: Sequence 1, Application US/09969671A
: Publication No. US20030036112A1
: GENERAL INFORMATION:
: APPLICANT: CHAPMAN, CONRAD G.
: APPLICANT: MURPHY, KAY
: APPLICANT: POWELL, DAVID J.
: APPLICANT: SMITH, TRUDI S.
: TITLE OF INVENTION: ASP2
: FILE REFERENCE: GH-70368-D1
: CURRENT APPLICATION NUMBER: US/09/969,671A
: CURRENT FILING DATE: 2001-10-03
: PRIOR APPLICATION NUMBER: UK 9701684.4
: PRIOR FILING DATE: 1997-01-28
: PRIOR APPLICATION NUMBER: 09/009,191
: PRIOR FILING DATE: 1998-01-20
: PRIOR APPLICATION NUMBER: 09/694,200
: PRIOR FILING DATE: 2000-10-23
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 2541
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unknown
: LOCATION: (2455)(2456)(2463)(2478)(2480)(2497)(2507)(2509)(2512)(2516)
: LOCATION: (2520)(2522)(2525)(2529)(2539)(2540)
: OTHER INFORMATION: Wherein n can be represented by a, c, t, or g
US-09-969-671A-1

Query Match 86.2%; Score 2024; DB 9; Length 2541;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2038; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 306 ATGGCCCAAGCCCTCGCCTGCTGCTGCTGATGGCGGGGAGTGCCTGGCCAC 365
Db 1 ATGGCCCAAGCCCTCGCCTGCTGCTGCTGATGGCGGGGAGTGCCTGGCCAC 60
QY 366 GGCACCCAGCAGCGCATCCCGCTGCCCTGCGCAGCGGCTGGGGGCGCCCTGGGG 425
Db 61 GGCACCCAGCAGCGCATCCCGCTGCCCTGCGCAGCGGCTGGGGGCGCCCTGGGG 120
QY 426 CTGCGGCTGCCCGGAGACCGAGAGAGCCCGAGAGCCCGGAGCGCCCGGAGGGAGCTTT 485
Db 121 CTGCGGCTGCCCGGAGACCGAGAGAGCCCGAGAGCCCGGAGCGCCCGGAGGGGAGCTTT 180
QY 486 GTGAGATGTGTGACAACTTGTAGGGCACTCGGGGAGGCGTACTACGTGGAGATCAC 545
Db 181 GTGAGATGTGTGACAACTTGTAGGGGCACTCGGGGAGGCGTACTACGTGGAGATGACC 240
QY 546 GTGGGAGCCCCCGCAGACGCTCAACATCTGTGTGATACAGCAGCAGTAACTTTGCA 605
Db 241 GTGGGAGCCCCCGCAGACGCTCAACATCTGTGTGATACAGCAGCAGTAACTTTGCA 300
QY 606 GTGGGTGCTGCCGCCACCCCTTCCTGCATCGCTACTACAGAGGCGAGCTGTCCAGACA 665
Db 301 GTGGGTGCTGCCGCCACCCCTTCCTGCATCGCTACTACCAGAGGCGAGCTGTCCAGACA 360
QY 666 TACCGGACCTCCGGAAGGTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAG 725
Db 361 TACCGGACCTCCGGAAGGTGTGTATGAGCCCTTACACCCAGGCAAGTGGGAAGGGAG 420
QY 726 CTGGGCACCGACCTGGTAAAGCATCCCCCATGGCCCAACGCTCACTGTGCGTGCCAACATT 785
```

```
|||||
Db 421 CTGGCACCACCTGGTAAAGCATCCCCCATGGCCCCAACGTCAGTGTGCGTGCCAACATT 480
QY 786 GCTGGCATCACTGAATCAGACAAGTTCTTTCATCAACGGCTCCCACTGGGAAGGATCCTG 845
Db 481 GCTGGCATCACTGAATCAGACAAGTTCTTTCATCAACGGCTCCCACTGGGAAGGATCCTG 540
QY 846 GGGCTGGCCTATGCTGAGATTGGCAGGCCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 905
Db 541 GGGCTGGCCTATGCTGAGATTGGCAGGCCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 600
QY 906 CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCGAGCTTTGTGTGCTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCGAGCTTTGTGTGCTGGCTTC 660
QY 966 CCCTCAACAGTCTGAAGTGTCTGGCCTCTGTGGAGGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCTCAACAGTCTGAAGTGTCTGGCCTCTGTGGAGGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCCATCCGCGGGAGTGTATTAT 1085
Db 721 GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCCATCCGCGGGAGTGTATTAT 780
QY 1086 GAGGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 1145
Db 781 GAGGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 840
QY 1146 TACAATATGACAAAGAGCATTTGGACAGTGGCAGCAGCAACCTTCGTTGCCCAAGAAA 1205
Db 841 TACAATATGACAAAGAGCATTTGGACAGTGGCAGCAGCAACCTTCGTTGCCCAAGAAA 900
QY 1206 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCTCTCCACGAGAGAGTTCCTGTAT 1265
Db 901 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCTCTCCACGAGAGAGTTCCTGTAT 960
QY 1266 GTTTCCTGGTAGGAGAGCAGTGTGTGTGTGCGAAGCAGCAGCACCCTTTGGAACATT 1325
Db 961 GTTTCCTGGTAGGAGAGCAGTGTGTGTGTGCGAAGCAGCAGCACCCTTTGGAACATT 1020
QY 1326 TTCCAGTCATCTCAGTCTACCTTAATGGGTGAGGTTACCAACAGTCTTCCTCGCATACC 1385
Db 1021 TTCCAGTCATCTCAGTCTACCTTAATGGGTGAGGTTACCAACAGTCTTCCTCGCATACC 1080
QY 1386 ATCTTCCTCCGAGCAATACCTCGCGCCAGTGGGAAGATGTGGCAGCTCCCAAGACAGCTGT 1445
Db 1081 ATCTTCCTCCGAGCAATACCTCGCGCCAGTGGGAAGATGTGGCAGCTCCCAAGACAGCTGT 1140
QY 1446 TACAAGTTTGGCATCTCAGAGTCATCCAGCGGCAGCTGTATGGAGCTGTATTATCATCGAG 1505
Db 1141 TACAAGTTTGGCATCTCAGAGTCATCCAGCGGCAGCTGTATTATGGAGCTGTATTATCATCGAG 1200
QY 1506 GGCTTCTACGTTGTCTTTGATCGGGGCGGAAAACGAATTTGGCTTGTGTCAGGCGCTTGC 1565
Db 1201 GGCTTCTACGTTGTCTTTGATCGGGGCGGAAAACGAATTTGGCTTGTGTCAGGCGCTTGC 1260
QY 1566 CATGTGCACAGTGTGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTGTACCTTTGGACATG 1625
Db 1261 CATGTGCACAGTGTGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTGTACCTTTGGACATG 1320
QY 1626 GAAGACTGTGGCTACAAATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
Db 1321 GAAGACTGTGGCTACAAATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380
QY 1686 GTCATGGCTGCCATCTGGGCCCTCTTTCATGCTGCCACTCTGCCCTCATGCTGTGTCAGTGG 1745
Db 1381 GTCATGGCTGCCATCTGGGCCCTCTTTCATGCTGCCACTCTGCCCTCATGCTGTGTCAGTGG 1440
QY 1746 CGCTGCCCTCGCTCGCTGGCCAGCAGCATGATGACTTTGCTGTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGCCCTCGCTCGCTGGCCAGCAGCATGATGACTTTGCTGTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGAGGAGGCCCATGGGCAAGAGATGAGATTCCCTT -GGACACACCTCCGCTGGTTC 1864
|||||
```


Db 1501 AAGTGAGAGGCCCCATGGGAGAAAAGATAGAGATTCCCTCGGGACCAACACCTCCGTGGTTC 1560
QY 1865 ACTTTGGTCAACAAGTAGGAGACACAGATGGCACTGTGGCCAGAGCAGCCTCAGGACCCCTC 1924
Db 1561 ACTTTGGTCAACAAGTAGGAGACACAGATGGCACTGTGGCCAGAGCAGCCTCAGGACCCCTC 1620
QY 1925 CCCACCCACCAATGCCCTCTGCTTGTATGGAGAAAGAAAGGCTGGCAAGGTTGCCA 1984
Db 1621 CCCACCCACCAATGCCCTCTGCTTGTATGGAGAAAGAAAGGCTGGCAAGGTTGCCA 1680
QY 1985 GGGACTGTACCTGTAGGAACAGAAAAGAGAAAGAAAGCACTCTGCTGGCGGAATAC 2044
Db 1681 GGGACTGTACCTGTAGGAACAGAAAAGAGAAAGAAAGCACTCTGCTGGCGGAATAC 1740
QY 2045 TCTTGGTCACTCAATTAAGTCGGGAAATCTGCTGTGAAACTTCAGCCCTGAACC 2104
Db 1741 TCTTGGTCACTCAATTAAGTCGGGAAATCTGCTGTGAAACTTCAGCCCTGAACC 1800
QY 2105 TTGTGCCACCATTCCTTTAAATTCCTCAACCCCAAGATTTCTTCTTTTCTTAGTTTCAGA 2164
Db 1801 TTGTGCCACCATTCCTTTAAATTCCTCAACCCCAAGATTTCTTCTTTTCTTAGTTTCAGA 1860
QY 2165 AGTACTGGCATCACAGCAGGTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAA 2224
Db 1861 AGTACTGGCATCACAGCAGGTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAA 1920
QY 2225 GAGACCAAGCTTGTTCCTGTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTAT 2284
Db 1921 GAGACCAAGCTTGTTCCTGTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTAT 1980
QY 2285 TTGCTTTAGACAGAGGACTGTATAAACAGCCTAACATTTGGTGCAAGATTGGCTCTTTG 2344
Db 1981 TTGCTTTAGACAGAGGACTGTATAAACAGCCTAACATTTGGTGCAAGATTGGCTCTTTG 2040
QY 2345 AATT 2348
Db 2041 AATT 2044

RESULT 8
US-09-795-903A-1
; Sequence 1, Application US/09795903A
; Patent No. US20020164760A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
; FILE REFERENCE: OMRF 179
; CURRENT APPLICATION NUMBER: US/09/795,903A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/604,608
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-903A-1

Query Match 85.3%; Score 2004; DB 9; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 GCGGAGTGTGCTGCTGCCACCGCACCCAGCATCCGGTTCGCCCTTCGCGAGCGC 404
Db 1 GCGGAGTGTGCTGCTGCCACCGCACCCAGCATCCGGTTCGCCCTTCGCGAGCGC 60
QY 405 CTGGGGGCGCCCTCGGGCTGCGGCTGCGGAGAGCCGAGAGAGCCGAGAG 464
Db 61 CTGGGGGCGCCCTCGGGCTGCGGCTGCGGAGAGCCGAGAGAGCCGAGAG 120
QY 465 CCGGCGGAGGCGGAGCTTTGTGGAGATGGTGACAACTTGAGGGCGAGT 524
Db 121 CCGGCGGAGGCGGAGCTTTGTGGAGATGGTGACAACTTGAGGGCGAGT 180
QY 525 GGCTACTACGTGGAGATGACCGTGGGCGAGCCCGCGAGAGCGCTCAACATCTCTGTGTGAT 584
Db 181 GGCTACTACGTGGAGATGACCGTGGGCGAGCCCGCGAGAGCGCTCAACATCTCTGTGTGAT 240
QY 585 ACAGGCGAGCTTAACCTTTGTCAGTGGTGTGCGCCCGCCACCGCTTCTGTCATCGCTTAC 644
Db 241 ACAGGCGAGCTTAACCTTTGTCAGTGGTGTGCGCCCGCCACCGCTTCTGTCATCGCTTAC 300
QY 645 CAGAGGAGCTTCCAGCACATACCGGACCTCCGGAAGGTGTGTATGTGCCCTTACAC 704
Db 301 CAGAGGAGCTTCCAGCACATACCGGACCTCCGGAAGGTGTGTATGTGCCCTTACAC 360
QY 705 CAGGCGAAGTGGGAAGGAGCTGGGCAACCGACTGTGTAAAGCATCCCGCCATGCCCCAAC 764
Db 361 CAGGCGAAGTGGGAAGGAGCTGGGCAACCGACTGTGTAAAGCATCCCGCCATGCCCCAAC 420
QY 765 GTCACTGTGCTGCCAACAATTTGTGTCATCACTGAATCAGACAAGTTCTTTCATCAACGCG 824
Db 421 GTCACTGTGCTGCCAACAATTTGTGTCATCACTGAATCAGACAAGTTCTTTCATCAACGCG 480
QY 825 TCCAACTCGGAAGGAGCTCCTGGGCTGGCTATGCTCAGATTTGCCAGGCTGACACATCC 884
Db 481 TCCAACTCGGAAGGAGCTCCTGGGCTGGCTATGCTCAGATTTGCCAGGCTGACACATCC 540
QY 885 CTGGAGCCTTTCTTTGACTCTCTGTGTAAGAGCACCCAGCTTCCCAAGCTCTTCTTCCCTG 944
Db 541 CTGGAGCCTTTCTTTGACTCTCTGTGTAAGAGCACCCAGCTTCCCAAGCTCTTCTTCCCTG 600
QY 945 CAGCTTTGTGCTGGCTTCCCTTCAACAGCTGTGAAGTGTGCTGCTGCGAGG 1004
Db 601 CAGCTTTGTGCTGGCTTCCCTTCAACAGCTGTGAAGTGTGCTGCTGCGAGG 660
QY 1005 AGCATGATCAATTGGAGGTATCGACCACTCGCTGTACACAGGAGCTCTCTGGTATACAC 1064
Db 661 AGCATGATCAATTGGAGGTATCGACCACTCGCTGTACACAGGAGCTCTCTGGTATACAC 720
QY 1065 ATCCGGCGGAGTGGTATTTATGAGGTGATCAATTGTGGGGTGGAGATCAATGGAGAGGAT 1124
Db 721 ATCCGGCGGAGTGGTATTTATGAGGTGATCAATTGTGGGGTGGAGATCAATGGAGAGGAT 780
QY 1125 CTGAAAATGGACTGCAAGGAGTACAATATGACAAGAGCATTTGGGAGCTGGGACCAACC 1184
Db 781 CTGAAAATGGACTGCAAGGAGTACAATATGACAAGAGCATTTGGGAGCTGGGACCAACC 840
QY 1185 AACCTTCTTCCCAAGAAAGTTTGAAGCTGCACTCAATCCATCAAGGAGCTTCC 1244
Db 841 AACCTTCTTCCCAAGAAAGTTTGAAGCTGCACTCAATCCATCAAGGAGCTTCC 900
QY 1245 TCCACGGAGAAAGTTCCCTGATGGTCTGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 1304
Db 901 TCCACGGAGAAAGTTCCCTGATGGTCTGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 960
QY 1305 GGCACCACTTGGAAACATTTTCCAGTCTATCTCACTCTACCTTAATGGGTGAGGTACC 1364
Db 961 GGCACCACTTGGAAACATTTTCCAGTCTATCTCACTCTACCTTAATGGGTGAGGTACC 1020
QY 1365 AACAGTCTTCCGATCACCCTTCCCGCAGCAATACCTGCGGCCAGTGGGAAGATGTG 1424
Db 1021 AACAGTCTTCCGATCACCCTTCCCGCAGCAATACCTGCGGCCAGTGGGAAGATGTG 1080

```
QY 1425 GCCACGTCACAGACGACTGTTACAAAGTTGGCATCTCACAGTCATCCACGGGCACTGTT 1484
|
|
|
Db 1081 GCCACGTCCCAAGACGACTGTTACAAAGTTGGCATCTCACAGTCATCCACGGGCACTGTT 1140
|
|
|
QY 1485 ATGGGAGCTGTTATCATGAGGGCTTCTAGCTTGCTTTGATCGGGCCGAAACAAAT 1544
|
|
|
Db 1141 ATGGGAGCTGTTATCATGAGGGCTTCTAGCTTGCTTTGATCGGGCCGAAACAAAT 1200
|
|
|
QY 1545 GGCCTTGCTGTCAGGCGCTTGCCATGTGCACATGTCCAGGACGCGCAGGGTGAAGGC 1604
|
|
|
Db 1201 GGCCTTGCTGTCAGGCGCTTGCCATGTGCACATGTCCAGGACGCGCAGGGTGAAGGC 1260
|
|
|
QY 1605 CCTTTTGTCACTTGGACATGGAAGACTGTGGCTACAACATCCACAGACAGATGAGTCA 1664
|
|
|
Db 1261 CCTTTTGTCACTTGGACATGGAAGACTGTGGCTACAACATCCACAGACAGATGAGTCA 1320
|
|
|
QY 1665 ACCCTCATGACCATAGCCTATGTCATGCGTGGCATCTGGCCCTCTTCATGCTGCCACTC 1724
|
|
|
Db 1321 ACCCTCATGACCATAGCCTATGTCATGCGTGGCATCTGGCCCTCTTCATGCTGCCACTC 1380
|
|
|
QY 1725 TGCTCATGCTGTGTCAGTGGCGCTGCCCTCGCTGCTGCGCCAGCAGCATGATGACTTT 1784
|
|
|
Db 1381 TGCTCATGCTGTGTCAGTGGCGCTGCCCTCGCTGCTGCGCCAGCAGCATGATGACTTT 1440
|
|
|
QY 1785 GCTGATGACATCTCCCTGCTGAAGTGAAGGAGGCCCCATGGGAGAGATAGAGATTCGCCCT 1844
|
|
|
Db 1441 GCTGATGACATCTCCCTGCTGAAGTGAAGGAGGCCCCATGGGAGAGATAGAGATTCGCCCT 1500
|
|
|
QY 1845 GGACACACCTCGTGGTTCACCTTTGGTCAAGTAGGAGACACATGGCAGCTGTGGC 1904
|
|
|
Db 1501 GGACACACCTCGTGGTTCACCTTTGGTCAAGTAGGAGACACATGGCAGCTGTGGC 1560
|
|
|
QY 1905 CAGAGCACCTCAGGACCCCTCCACCCACCAATGCTGCTGCTGATGGAGAGGAAAA 1964
|
|
|
Db 1561 CAGAGCACCTCAGGACCCCTCCACCCACCAATGCTGCTGCTGATGGAGAGGAAAA 1620
|
|
|
QY 1965 GCGTGCAAGTGGGTTCCAGGAGCTGTATCTGTAGGAACAGAAAAGAGAAAGAAAG 2024
|
|
|
Db 1621 GCGTGCAAGTGGGTTCCAGGAGCTGTATCTGTAGGAACAGAAAAGAGAAAGAAAG 1680
|
|
|
QY 2025 CACTGCTGCGGGGGAATCTTGTGTCACCTCAAAATTTAAGTCGGGAAATCTGCTGCT 2084
|
|
|
Db 1681 CACTGCTGCGGGGGAATCTTGTGTCACCTCAAAATTTAAGTCGGGAAATCTGCTGCT 1740
|
|
|
QY 2085 TGAACCTCAGGCGCTGAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT 2144
|
|
|
Db 1741 TGAACCTCAGGCGCTGAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT 1800
|
|
|
QY 2145 CTTCTTTTCTTAGTTTTCAGAACTGAGGATCAGCGAGTTACCTTGGCGTGTGTCCTC 2204
|
|
|
Db 1801 CTTCTTTTCTTAGTTTTCAGAACTGAGGATCAGCGAGTTACCTTGGCGTGTGTCCTC 1860
|
|
|
QY 2205 TGTGGTACCTTGGCAGAGAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGA 2264
|
|
|
Db 1861 TGTGGTACCTTGGCAGAGAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGA 1920
|
|
|
QY 2265 GAGGATGCACAGTTTCTGCTATTTCTTTAGAGACAGGAGCTGTATAACAAAGCCTAACATT 2324
|
|
|
Db 1921 GAGGATGCACAGTTTCTGCTATTTCTTTAGAGACAGGAGCTGTATAACAAAGCCTAACATT 1980
|
|
|
QY 2325 GGTGCAAGATTGCCCTCTTGAATT 2348
|
|
|
Db 1981 GGTGCAAGATTGCCCTCTTGAATT 2004
|
|
|
```

RESULT 9
US-09-796-264-1
; Sequence 1, Application US/09796264
; Patent No. US20020049303A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald

```
; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
; FILE REFERENCE: OMRF 179  
; CURRENT APPLICATION NUMBER: US/09/796,264  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 09/604,608  
; PRIOR FILING DATE: 2000-06-27  
; PRIOR APPLICATION NUMBER: 60/168,060  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: 60/177,836  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/178,368  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: 60/210,292  
; PRIOR FILING DATE: 2000-06-08  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3252  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-264-1  
  
Query Match 85.3%; Score 2004; DB 10; Length 3252;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 345 GCGGGAGTGCTGCTGCCACAGCGCACCCAGCAGCGCATCGCGCTGCCCTGCCAGCGGC 404  
|  
|  
|  
Db 1 GCGGGAGTGCTGCTGCCACAGCGCACCCAGCAGCGCATCGCGCTGCCCTGCCAGCGGC 60  
|  
|  
|  
QY 405 CTGGGGGGCGCCCGCTGGGGCTGCGGCTGCCCGGGAGACCGACGAGAGCGCCGAGGAG 464  
|  
|  
|  
Db 61 CTGGGGGGCGCCCGCTGGGGCTGCGGCTGCCCGGGAGACCGACGAGAGCGCCGAGGAG 120  
|  
|  
|  
QY 465 CCGCGCGGAGGGGCGAGCTTTTGTGGAGATGGTGGACAACTGAGGGGCAAGTCGGGGCAG 524  
|  
|  
|  
Db 121 CCGCGCGGAGGGGCGAGCTTTTGTGGAGATGGTGGACAACTGAGGGGCAAGTCGGGGCAG 180  
|  
|  
|  
QY 525 GGCTACTACGTGAGATGACCTGGGCGAGCGCCCGCAGAGCGCTCAACATCTCTGGTGGAT 584  
|  
|  
|  
Db 181 GGCTACTACGTGAGATGACCTGGGCGAGCGCCCGCAGAGCGCTCAACATCTCTGGTGGAT 240  
|  
|  
|  
QY 585 ACAGGCGAGTAGTAACCTTTGCAGTGGGTGCTGCGCCCGCCACCCCTTCTCGCATCGCTACTAC 644  
|  
|  
|  
Db 241 ACAGGCGAGTAGTAACCTTTGCAGTGGGTGCTGCGCCCGCCACCCCTTCTCGCATCGCTACTAC 300  
|  
|  
|  
QY 645 CAGAGCGAGCTGTCAGCACATACCGGGACCTCCGGAGGGTGTGTATGTGCCCTACACC 704  
|  
|  
|  
Db 301 CAGAGCGAGCTGTCAGCACATACCGGGACCTCCGGAGGGTGTGTATGTGCCCTACACC 360  
|  
|  
|  
QY 705 CAGGGCAAGTGGGAAGGGAGCTGGGCGACCGACCTGTAAGCATCCCGCATGGCCCCAAC 764  
|  
|  
|  
Db 361 CAGGGCAAGTGGGAAGGGAGCTGGGCGACCGACCTGTAAGCATCCCGCATGGCCCCAAC 420  
|  
|  
|  
QY 765 GTCACCTGTCGGTGCCCAACATTTGCTGCTACCTCACTGAATCAGACAAAGTCTTTCATCAACGGC 824  
|  
|  
|  
Db 421 GTCACCTGTCGGTGCCCAACATTTGCTGCTACCTCACTGAATCAGACAAAGTCTTTCATCAACGGC 480  
|  
|  
|  
QY 825 TCCAACCTGGGAAGGCAATTCCTGGGGCTGGCCTATGCTGAGATGGCAGGCTGAGGACTCC 884  
|  
|  
|  
Db 481 TCCAACCTGGGAAGGCAATTCCTGGGGCTGGCCTATGCTGAGATGGCAGGCTGAGGACTCC 540  
|  
|  
|  
QY 885 CTGGAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCTCCCTG 944  
|  
|  
|  
Db 541 CTGGAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCTCCCTG 600  
|  
|  
|  
QY 945 CAGCTTTGTGCTGCTGGCTTCCCGCTCAACAGCTCTCAAGTGTGGCTCTGTGGGAGGG 1004  
|  
|  
|  
Db 601 CAGCTTTGTGCTGCTGGCTTCCCGCTCAACAGCTCTCAAGTGTGGCTCTGTGGGAGGG 660  
|  
|  
|  
QY 1005 AGCATGATCATTTGGAGGTATCGACCACTCGCTGTATACAGGAGCTCTCTGTATACACC 1064  
|  
|  
|
```


[illegible]

```

RESULT 11
US-09-794-927-5
; Sequence 5, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794, 927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416, 901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155, 493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404, 133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101, 594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1977

```


Db 1966 A 1966

RESULT 12

```

US-09-795-847-5
; Sequence 5, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795,847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatencIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-847-5

```

Query Match 80.1%; Score 1881; DB 10; Length 1977;
Best Local Similarity 96.3%; Pred. NO. 0;
Matches 1966; Conservative 0; Mismatches 0; Indels 75; Gaps 1;

Qy	306	ATGGCCCAAGCCCTGCCCTGGCTCCTCTGCTGTGATGGGCGGGAGTGTGCTGTGCCAC	365
Db	1	ATGGCCCAAGCCCTGCCCTGGCTCCTCTGCTGTGATGGGCGGGAGTGTGCTGTGCCAC	60
Qy	366	GGACCCAGCAGCAGCATCCGGCTGCCCTGCGCAGCGGCTGGGGGGGCCGCCCCCTGGGG	425
Db	61	GGACCCAGCAGCGCATCCGGCTGCCCTGCGCAGCGGCTGGGGGGGCCGCCCTGGGG	120
Qy	426	CTGGGGTGTCCCGGGAGACCGACGAAGAGCCGAGAGGCCGGCGAGGGGACGCTTTT	485
Db	121	CTGGGTGTCCCGGGAGACCGACGAAGAGCCCGAGAGCCCGCCGGAGGGGACGCTTT	180
Qy	486	GTGGAGATGTGGACAACTTGAGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC	545
Db	181	GTGGAGATGTGGACAACTTGAGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC	240
Qy	546	GTGGGACGCCCCCGCAGACGCTCAACATCTCGTGATACAGGCAGCAGTAACTTTGCA	605
Db	241	GTGGGACGCCCCCGCAGACGCTCAACATCTCGTGATACAGGCAGCAGTAACTTTGCA	300
Qy	606	GTGGGTGCTGGCCCCCACCCCTTCCCTGCGATCTACCAGAGGCGAGTGTCCAGCACA	665
Db	301	GTGGGTGCTGGCCCCCACCCCTTCCCTGCGATCTACCAGAGGCGAGTGTCCAGCACA	360
Qy	666	TACCGGACCTCCGGGAAGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG	725
Db	361	TACCGGACCTCCGGGAAGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG	420
Qy	726	CTGGGCACGACCTGTGAAGCATCCCCCATGGCCCCAACGTCACTGTGGGTGCCACACTT	785
Db	421	CTGGGCACGACCTGTGAAGCATCCCCCATGGCCCCAACGTCACTGTGGGTGCCACACTT	480

Qy	786	GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCGTG	845
Db	481	GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCGTG	540
Qy	846	GGCTGSCCTATGCTGAGATTGCCAGSCTTGACGACTCCCTGGAGCCTTTCTTTGACTCT	905
Db	541	GGCTGSCCTATGCTGAGATTGCCAGSCTTGACGACTCCCTGGAGCCTTTCTTTGACTCT	566
Qy	906	CTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCCTCCCTGCAGCTTTGTGGTGTGCTGGCTTC	965
Db	567	CTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCCTCCCTGCAGCTTTGTGGTGTGCTGGCTTC	585
Qy	966	CCCTCAACCGAGTCTGAAGTGTGGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC	1021
Db	586	CCCTCAACCGAGTCTGAAGTGTGGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC	645
Qy	1026	GACCACTCGCTGTACACAGCAGCTCTGGTATACACCCATCCGGCGGAGTGGTATTAT	1081
Db	646	GACCACTCGCTGTACACAGCAGCTCTGGTATACACCCATCCGGCGGAGTGGTATTAT	705
Qy	1086	GAGTGATCATTTGTGGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	1141
Db	706	GAGTGATCATTTGTGGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	765
Qy	1146	TACAACATATGACAAGAGCATTTGGGAGTGGCACCACCAACCTTCGTTTGGCCAAAGAA	1201
Db	766	TACAACATATGACAAGAGCATTTGGGAGTGGCACCACCAACCTTCGTTTGGCCAAAGAA	825
Qy	1206	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCAGCCTCTCCACGAGAGTTCCCTGAT	1261
Db	826	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCAGCCTCTCCACGAGAGTTCCCTGAT	885
Qy	1266	GGTTTCTGGCTAGGAGCAGCTGGTGTGCTGGCAAGCAGCACACCCCTTGGGAACATT	1321
Db	886	GGTTTCTGGCTAGGAGCAGCTGGTGTGCTGGCAAGCAGCACACCCCTTGGGAACATT	945
Qy	1326	TTCCAGTCATCTCACTCTACCTTAATGGGTGAGTTTACCAACAGTCTTCCGCGATCAC	1381
Db	946	TTCCAGTCATCTCACTCTACCTTAATGGGTGAGTTTACCAACAGTCTTCCGCGATCAC	1001
Qy	1386	ATCCTTCCGCGAGCAATACCTTGGCGGAGTGGGAAGATGTGGCCACGTCCCAAGACGACTGT	1441
Db	1006	ATCCTTCCGCGAGCAATACCTTGGCGGAGTGGGAAGATGTGGCCACGTCCCAAGACGACTGT	1061
Qy	1446	TACAAGTTTGGCATCTCACAGTCATCCACGGGCACGTGTTATGGAGCTGTATTATCATGGAG	1501
Db	1066	TACAAGTTTGGCATCTCACAGTCATCCACGGGCACGTGTTATGGAGCTGTATTATCATGGAG	1121
Qy	1506	GGCTTTCACGTTGTCTTTGATCGGGGCCGAAAACGAATTTGGCTTTGCTGTCAAGCGCTGC	1561
Db	1126	GGCTTTCACGTTGTCTTTGATCGGGGCCGAAAACGAATTTGGCTTTGCTGTCAAGCGCTGC	1181
Qy	1566	CATGTGCACATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTCACTTTGGACATG	1621
Db	1186	CATGTGCACATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTCACTTTGGACATG	1241
Qy	1626	GAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT	1681
Db	1246	GAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT	1301
Qy	1686	GTATGCTGCCATCTCGCGCCCTTTCATGCTGCCACTTGCCTCATGTTGTTGCTCAAGTGG	1741
Db	1306	GTATGCTGCCATCTCGCGCCCTTTCATGCTGCCACTTGCCTCATGTTGTTGCTCAAGTGG	1361
Qy	1746	CGCTGCTCGCTCGCTTGGCGGACGACCATGATGACTTTGCTGATGACATCTCCCTGCTG	1801
Db	1366	CGCTGCTCGCTCGCTTGGCGGACGACCATGATGACTTTGCTGATGACATCTCCCTGCTG	1421
Qy	1806	AAGTGGAGGCCCATCGGCGAGAGATAGAGATTTCCCTTGGACACACACCTTCGTTGTTCA	1861
Db	1426	AAGTGGAGGCCCATCGGCGAGAGATAGAGATTTCCCTTGGACACACACCTTCGTTGTTCA	1481
Qy	1866	CTTTGGTCAAAAGTATGAGGACACAGATGGCAGCTGTGGCCAGAGCAGCTTCAGGACCCCTCC	1921

1486 CTTTGGTCAACAAGTAGGAGACACAGATGGACCTGTGGCCAGAGACACCTCAGACCCCTCC 1545
1926 CCACCCACCAAAATGCTTCTGCTTTCATGGAGAGGAAAGGCTGGCAAGTGGGTTCAG 1985
1546 CCACCCACCAAAATGCTTCTGCTTTCATGGAGAGGAAAGGCTGGCAAGTGGGTTCAG 1605
1986 GGACTGTACCTGTAGGAACAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2045
1606 GGACTGTACCTGTAGGAACAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1665
2046 CTTTGGTCAACCTCAAAATTAAGTGGGAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 2105
1666 CTTTGGTCAACCTCAAAATTAAGTGGGAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 1725
2106 TTGTGCCACCAATTCCTTTAAATTTCTCAACCCCAAGATATTCTTCTTTTCTTTAGTTTCAGAA 2165
1726 TTGTGCCACCAATTCCTTTAAATTTCTCAACCCCAAGATATTCTTCTTTTCTTTAGTTTCAGAA 1785
2166 GTACTGGCATCACACGAGGTTTACCTTGGCGGTGTGTCTTCCCTGTGGTACCCCTGGCAGAGAAG 2225
1786 GTACTGGCATCACACGAGGTTTACCTTGGCGGTGTGTCTTCCCTGTGGTACCCCTGGCAGAGAAG 1845
2226 AGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGGCTATT 2285
1846 AGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGGCTATT 1905
2286 TGCTTTAGAGACAGGGAGCTGTATAAACAAGCCTTAACATTGGTGCAAGATTGGCTCTTTGA 2345
1906 TGCTTTAGAGACAGGGAGCTGTATAAACAAGCCTTAACATTGGTGCAAGATTGGCTCTTTGA 1965
2346 A 2346
1966 A 1966

RESULT 13
US-09-794-743-5
; Sequence 5, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Blenkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-794-743-5

Query Match 80.1%; Score 1881; DB 10; Length 1977;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1966; Conservative 0; Mismatches 0; Indels 75; Gaps 1;

306 ATGCCCCAAGCCCTGCCCTGGCTCCTGTGTGTATGGCGCGGAGTGTCTGCCCTGCCAC 365
1 ATGCCCCAAGCCCTGCCCTGGCTCCTGTGTGTATGGCGCGGAGTGTCTGCCCTGCCAC 60
366 GGCACCCAGCAGCAGATCCCGCTGCCCTGTGCGAGAGCGGCTGGGGGGCGCCCTGGGG 425
61 GGCACCCAGCAGCAGATCCCGCTGCCCTGTGCGAGAGCGGCTGGGGGGCGCCCTGGGG 120
426 CTGCGGCTGCCCGGGAGACCGACGAAGAGCCCGAGAGCCCGAGAGCCCGAGAGCGGCTTT 485
121 CTGCGGCTGCCCGGGAGACCGACGAAGAGCCCGAGAGCCCGAGAGCGGCTTT 180
486 GTGAGATGTGTGACAACTCTGAGGGCAAGTGGGGCAGGGCTTACTACGTGGAGATGACC 545
181 GTGAGATGTGTGACAACTCTGAGGGCAAGTGGGGCAGGGCTTACTACGTGGAGATGACC 240
546 GTGGGAGCCCGCCGAGACGCTCAACATCTCTGGTGGATACAGGAGCAGTAACTTTGCA 605
241 GTGGGAGCCCGCCGAGACGCTCAACATCTCTGGTGGATACAGGAGCAGTAACTTTGCA 300
606 GTGGGAGCCCGCCGAGACGCTCAACATCTCTGGTGGATACAGGAGCAGTAACTTTGCA 665
301 GTGGGAGCCCGCCGAGACGCTCAACATCTCTGGTGGATACAGGAGCAGTAACTTTGCA 360
666 TACCGGAGCCCTCCGGAAGGGTGTGTATGTGCCCTTACACCCAGGGCAAGTGGGAAGGGGAG 725
361 TACCGGAGCCCTCCGGAAGGGTGTGTATGTGCCCTTACACCCAGGGCAAGTGGGAAGGGGAG 420
726 CTGGGAGCCGAGCTGTGTAAGCATCCCGCATGGCCCAAGCTACTGTGCGGTGCCAACAATT 785
421 CTGGGAGCCGAGCTGTGTAAGCATCCCGCATGGCCCAAGCTACTGTGCGGTGCCAACAATT 480
786 GCTGCCATCACTGAATCAGACAAGTTCTTCAACACGGCTTCCAACTGGGAAGGATCCTG 845
481 GCTGCCATCACTGAATCAGACAAGTTCTTCAACACGGCTTCCAACTGGGAAGGATCCTG 540
846 GGGCTGCCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGGCTTCTTTTGACTCT 905
541 GGGCTGCCCTATGCTGAGATTGCCAG-----GCTTTGTGGTGGCTTTC 566
906 CTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGGCTTTC 965
567 -----GCTTTGTGGTGGCTTTC 585
966 CCCTCAACCAAGTCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTGGAGGTATC 1025
586 CCCTCAACCAAGTCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTGGAGGTATC 645
1026 GACCACTCGCTGTACACAGGAGCTCTGTGGTATACACCATCCCGGGGAGTGGTATAT 1085
646 GACCACTCGCTGTACACAGGAGCTCTGTGGTATACACCATCCCGGGGAGTGGTATAT 705
1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 1145
706 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 765
1146 TACAACCTATGACAAGAGCATTTGTGGACAGTGGCACCACCAACCTTCTTTGCCCAAGAAA 1205
766 TACAACCTATGACAAGAGCATTTGTGGACAGTGGCACCACCAACCTTCTTTGCCCAAGAAA 825
1206 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGAGGCTTCTTCCACGGAGAAGTTCCCTGAT 1265
826 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGAGGAGCTTCTTCCACGGAGAAGTTCCCTGAT 885
1266 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAGCAGGAGCACCACCTTGGAAACATT 1325
886 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAGCAGGAGCACCACCTTGGAAACATT 945
1326 TTCCCACTCATCTCACTTACCTAATGGGTGAGTTTACCAACCAAGCTTCTTCCCGCATACC 1385
946 TTCCCACTCATCTCACTTACCTAATGGGTGAGTTTACCAACCAAGCTTCTTCCCGCATACC 1005

QY 906 CTGTTAAAGCAGACCCACCTTCCCAACCTTCTCCCTGCAGCTTTTGTGGTGGCTTC 965
Db 567 -----GCTTTGGTGGCTGGCTTC 585
QY 966 CCCCTCAACCAAGTCTGAAGTGTGGCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC 1025
Db 586 CCCCTCAACCAAGTCTGAAGTGTGGCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC 645
QY 1026 GACCACCTCGCTGTACACAGGCACTCTCTGTATACACCCATCCGGCGGAGTGTATAT 1085
Db 646 GACCACCTCGCTGTACACAGGCACTCTCTGTATACACCCATCCGGCGGAGTGTATAT 705
QY 1086 GAGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAAATGGAGTGAAGGAG 1145
Db 706 GAGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAAATGGAGTGAAGGAG 765
QY 1146 TACAACCTATGACAGAGATTTGGACAGTGGACACCAACCTTCTGTTGCCCAAGAAA 1205
Db 766 TACAACCTATGACAGAGATTTGGACAGTGGACACCAACCTTCTGTTGCCCAAGAAA 825
QY 1206 GTGTTTGAAGCTGCAGTCAATCCATCAAGGAGCCTCTCCAGGAGAGTTCCTCGAT 1265
Db 826 GTGTTTGAAGCTGCAGTCAATCCATCAAGGAGCCTCTCCAGGAGAGTTCCTCGAT 885
QY 1266 GTTCTGCTAGGAGAGCAGTGGTGTGCTGGCAGAGCAGCCACCTTGGAAACAT 1325
Db 886 GTTCTGCTAGGAGAGCAGTGGTGTGCTGGCAGAGCAGCCACCTTGGAAACAT 945
QY 1326 TTCCCACTCATCTCACTTACCTAATGGTGAGTTTACCAACCAAGTCTTCCCGCATACC 1385
Db 946 TTCCCACTCATCTCACTTACCTAATGGTGAGTTTACCAACCAAGTCTTCCCGCATACC 1005
QY 1386 ATCTTCCGAGCAATACCTGCGGCGAGTGAAGATGTGGCCACGTCCTCCCAAGACGACTGT 1445
Db 1006 ATCTTCCGAGCAATACCTGCGGCGAGTGAAGATGTGGCCACGTCCTCCCAAGACGACTGT 1065
QY 1446 TACAAGTTTGGCATCTCAGTCACTACGAGGAGTGTATGGAGGTGTATATCATGAG 1505
Db 1066 TACAAGTTTGGCATCTCAGTCACTACGAGGAGTGTATGGAGGTGTATATCATGAG 1125
QY 1506 GGCTTCTACGTTGCTTTGATCGGCGCCGAGAAACGAATTTGGCTGTGTCAGCGCTTC 1565
Db 1126 GGCTTCTACGTTGCTTTGATCGGCGCCGAGAAACGAATTTGGCTGTGTCAGCGCTTC 1185
QY 1566 CATGTGCACGATGAGTTTCAGGACGCGAGCGGTGAAGGCGCTTTTGTACCTTTGGACATG 1625
Db 1186 CATGTGCACGATGAGTTTCAGGACGCGAGCGGTGAAGGCGCTTTTGTACCTTTGGACATG 1245
QY 1626 GAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1685
Db 1246 GAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1305
QY 1686 GTCATGCTGCGCTTCCGCTGCGCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1745
Db 1306 GTCATGCTGCGCTTCCGCTGCGCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1365
QY 1746 CGCTGCTGCGCTTCCGCTGCGCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1805
Db 1366 CGCTGCTGCGCTTCCGCTGCGCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1425
QY 1806 AAGTGAGGAGCCCATGGGCAAGATAGAGATTTCCCTGGACCAACCTTCCGTTGGTTCA 1865
Db 1426 AAGTGAGGAGCCCATGGGCAAGATAGAGATTTCCCTGGACCAACCTTCCGTTGGTTCA 1485
QY 1866 CTTTGTGTACAAGTGAAGACACAGATGCACTCTGTGGCCAGAGCAGCCTCAGGACCTCC 1925
Db 1486 CTTTGTGTACAAGTGAAGACACAGATGCACTCTGTGGCCAGAGCAGCCTCAGGACCTCC 1545
QY 1926 CCACCCACCAATGCTTCTGCTTGTATGGAGAGGAAAGGCTGGCAAGGTGGGTTCCAG 1985
Db 1546 CCACCCACCAATGCTTCTGCTTGTATGGAGAGGAAAGGCTGGCAAGGTGGGTTCCAG 1605

QY 1986 GGACTGTACCTGTAGGAAACAGAAAAGAGAAAGCACTCTGCTGGGGGAATACT 2045
Db 1606 GGACTGTACCTGTAGGAAACAGAAAAGAGAAAGCACTCTGCTGGGGGAATACT 1665
QY 2046 CTTGGTCACTCAAAATTTAAGTCGGGAAATTCCTGCTCTTGAACCTTCAGCCCTGAACCT 2105
Db 1666 CTTGGTCACTCAAAATTTAAGTCGGGAAATTCCTGCTCTTGAACCTTCAGCCCTGAACCT 1725
QY 2106 TTGTCACCACTTCCCTTAAATTTCTCAACCCAAAGTATTTCTTCTTCTAGTTTCAGAA 2165
Db 1726 TTGTCACCACTTCCCTTAAATTTCTCAACCCAAAGTATTTCTTCTTCTAGTTTCAGAA 1785
QY 2166 GTACTGGCATCACACGAGGTTACCTTGGCTGTGCTCCCTGTGTACCTTGGCAGAGAG 2225
Db 1786 GTACTGGCATCACACGAGGTTACCTTGGCTGTGCTCCCTGTGTACCTTGGCAGAGAG 1845
QY 2226 AGACCAAGCTTGTTCCTGCTGGCCAAAGTCTAGTAGGAGAGATGCAAGTTTCTCTAT 2285
Db 1846 AGACCAAGCTTGTTCCTGCTGGCCAAAGTCTAGTAGGAGAGATGCAAGTTTCTCTAT 1905
QY 2286 TGCTTTAGACAGAGGAGTGTATAAACAAGCCTTAACATTTGTCACCAAGATTCCTCTTCA 2345
Db 1906 TGCTTTAGACAGAGGAGTGTATAAACAAGCCTTAACATTTGTCACCAAGATTCCTCTTCA 1965
QY 2346 A 2346
Db 1966 A 1966
RESULT 15
US-09-794-925-5
; Sequence 5, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AN
; FILE REFERENCE: 28341/6280HI
; CURRENT APPLICATION NUMBER: US/09/794,925
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-794-925-5
Query Match 80.1%; Score 1881; DB 10; Length 1977;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1966; Conservative 0; Mismatches 0; Indels 75; Gaps 1;
QY 306 ATGCCCCAAGCCCTGCCCCCTGGCTTCTGCTGTGTGATGGCGCGGAGTGTGCTGCTGCCAC 365
Db 1 ATGCCCCAAGCCCTGCCCCCTGGCTTCTGCTGTGTGATGGCGCGGAGTGTGCTGCTGCCAC 60
QY 366 GGCAACCCAGCAGCAGCATCCGCTGCCCTGGCAGGCGCTGGGGGGCGCCCTCGGG 425
Db 61 GGCAACCCAGCAGCAGCATCCGCTGCCCTGGCAGGCGCTGGGGGGCGCCCTCGGG 120

```
QY 426 CTGCGCTGCCCGGAGACCGAGAGCCGAGAGCCGCGGAGGGGCGAGCTTT 485
|
|
|
Db 121 CTGCGCTGCCCGGAGACCGAGAGCCGAGAGCCGCGGAGGGGCGAGCTTT 180
|
|
|
QY 486 GTGAGATGGTGACAACTGAGGGGAAGTCGGGCGAGGCTACTACGTGGAGATGACC 545
|
|
|
Db 181 GTGAGATGGTGACAACTGAGGGGAAGTCGGGCGAGGCTACTACGTGGAGATGACC 240
|
|
|
QY 546 GTGGGAGCCCCCGAGAGCGCTACACATCTGTGTGATACAGGACGACGTAACCTTGCA 605
|
|
|
Db 241 GTGGGAGCCCCCGAGAGCGCTACACATCTGTGTGATACAGGACGACGTAACCTTGCA 300
|
|
|
QY 606 GTGGGTGCTGCCCCCCACCCCTTCCTGCACTGCTACTACAGAGGCAAGCTGTCCAGCACA 665
|
|
|
Db 301 GTGGGTGCTGCCCCCCACCCCTTCCTGCACTGCTACTACAGAGGCAAGCTGTCCAGCACA 360
|
|
|
QY 666 TACCGGACCTCGGGAAGGTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG 725
|
|
|
Db 361 TACCGGACCTCGGGAAGGTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG 420
|
|
|
QY 726 CTGGGACCGACCTGGTAAAGCATCCCCATGGCCCAACCTCACTGTGCTGCCAAGCAT 785
|
|
|
Db 421 CTGGGACCGACCTGGTAAAGCATCCCCATGGCCCAACCTCACTGTGCTGCCAAGCAT 480
|
|
|
QY 786 GCTGCCATCACTGAATCAGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG 845
|
|
|
Db 481 GCTGCCATCACTGAATCAGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG 540
|
|
|
QY 846 GGCTGGCTATGCTGAGATGGCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 905
|
|
|
Db 541 GGCTGGCTATGCTGAGATGGCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 566
|
|
|
QY 906 CTGGTAAAGCAGACCCACGTTCCCAACCTCTCTCCCTGCAGCTTGTGCTGGCTGGCTTC 965
|
|
|
Db 567 -----GCTTTGGTGGCTGGCTTC 585
|
|
|
QY 966 CCCCTCAACAGCTGAAGTGTGGCTCTGTGGGAGGAGCATGATCATTTGGAGGTATC 1025
|
|
|
Db 586 CCCCTCAACAGCTGAAGTGTGGCTCTGTGGGAGGAGCATGATCATTTGGAGGTATC 645
|
|
|
QY 1026 GACCACTCGCTACACAGGCGAGTCTGTGTATACACCCATCCGCGGGGAGTGGTATTAT 1085
|
|
|
Db 646 GACCACTCGCTACACAGGCGAGTCTGTGTATACACCCATCCGCGGGGAGTGGTATTAT 705
|
|
|
QY 1086 GAGTGATCATCTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 1145
|
|
|
Db 706 GAGTGATCATCTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 765
|
|
|
QY 1146 TACAACCTATGACAAGAGCATTTGGACAGTGGCACCAACCAACCTTCGTTTTGCCCAAGAAA 1205
|
|
|
Db 766 TACAACCTATGACAAGAGCATTTGGACAGTGGCACCAACCAACCTTCGTTTTGCCCAAGAAA 825
|
|
|
QY 1206 GTGTTTGAAGCTGAGTCAATCCATCAAGGCAAGCTTCCTCCAGGAGAAAGTTCCCTGAT 1265
|
|
|
Db 826 GTGTTTGAAGCTGAGTCAATCCATCAAGGCAAGCTTCCTCCAGGAGAAAGTTCCCTGAT 885
|
|
|
QY 1266 GGTTCCTGGCTAGGAGAGCTGGTGTGCTGGCAAGCAGCACCACCCCTTGGAACATT 1325
|
|
|
Db 886 GGTTCCTGGCTAGGAGAGCTGGTGTGCTGGCAAGCAGCACCACCCCTTGGAACATT 945
|
|
|
QY 1326 TTCCCAAGTCACTCACTTACCTTAATGGGTGAGGTTACCAACCAAGTCTTTCCCGCATCACC 1385
|
|
|
Db 946 TTCCCAAGTCACTCACTTACCTTAATGGGTGAGGTTACCAACCAAGTCTTTCCCGCATCACC 1005
|
|
|
QY 1386 ATCTTTCCGAGCAATACCTCGCGGCGAGTGGAAAGATGTGCCACGTCCTCCCAAGACGACTGT 1445
|
|
|
Db 1006 ATCTTTCCGAGCAATACCTCGCGGCGAGTGGAAAGATGTGCCACGTCCTCCCAAGACGACTGT 1065
|
|
|
QY 1446 TACAAGTTTCCCATCTCACAGTCACTCCACGGGCATCTTATGGAGCTGTATCATGGAG 1505
|
|
|
Db 1066 TACAAGTTTCCCATCTCACAGTCACTCCACGGGCATCTTATGGAGCTGTATCATGGAG 1125
|
|
|
```

```
QY 1506 GGCTTCTACGTTGTCTTTGATCGGGCCCCGAAACGAATTTGGCTTTGCTGTACGGCTTGC 1565
|
|
|
Db 1126 GGCTTCTACGTTGTCTTTGATCGGGCCCCGAAACGAATTTGGCTTTGCTGTACGGCTTGC 1185
|
|
|
QY 1566 CATGTGCACGATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1625
|
|
|
Db 1186 CATGTGCACGATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1245
|
|
|
QY 1626 GAAGACTGTGGCTACAACTTCCACACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
|
|
|
Db 1246 GAAGACTGTGGCTACAACTTCCACACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1305
|
|
|
QY 1686 GTCATGGCTGCCATCTGCGCCCTTTCATGCTGCACTCTGCCCTCATGCTGTGTCAGTGG 1745
|
|
|
Db 1306 GTCATGGCTGCCATCTGCGCCCTTTCATGCTGCACTCTGCCCTCATGCTGTGTCAGTGG 1365
|
|
|
QY 1746 CGCTGCCCTCGCTGCTGCGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1805
|
|
|
Db 1366 CGCTGCCCTCGCTGCTGCGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1425
|
|
|
QY 1806 AAGTGAGGAGGCCCATGGGAGAGATAGAGATTCCTCTGGACCACACCTCCCGTGGTTCA 1865
|
|
|
Db 1426 AAGTGAGGAGGCCCATGGGAGAGATAGAGATTCCTCTGGACCACACCTCCCGTGGTTCA 1485
|
|
|
QY 1866 CTTTGGTCAACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGACCTCAGGACCTCC 1925
|
|
|
Db 1486 CTTTGGTCAACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGACCTCAGGACCTCC 1545
|
|
|
QY 1926 CCACCCACCAATGCTCTGCTTTGATGGAGAGGAAAGCTGGCAAGTGGGTTCAG 1985
|
|
|
Db 1546 CCACCCACCAATGCTCTGCTTTGATGGAGAGGAAAGCTGGCAAGTGGGTTCAG 1605
|
|
|
QY 1986 GGACTGTACTGTAGGAAACAGAAAGAGAAAGCACTCTGCTGGCGGGAATACT 2045
|
|
|
Db 1606 GGACTGTACTGTAGGAAACAGAAAGAGAAAGCACTCTGCTGGCGGGAATACT 1665
|
|
|
QY 2046 CTTTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTCTTTGAAACTTTCAGTTCAGAA 2105
|
|
|
Db 1666 CTTTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTCTTTGAAACTTTCAGTTCAGAA 1725
|
|
|
QY 2106 TTGTCACGATTCCTTTAAATTCCTCAACCCCAAGATATCTCTTTTAGTTCAGAA 2165
|
|
|
Db 1726 TTGTCACGATTCCTTTAAATTCCTCAACCCCAAGATATCTCTTTTAGTTCAGAA 1785
|
|
|
QY 2166 GTACTGGCATCACACGAGGTTACCTTTGGCGTGTGCTGTGCTACCCCTGGCAGAGAAG 2225
|
|
|
Db 1786 GTACTGGCATCACACGAGGTTACCTTTGGCGTGTGCTGTGCTACCCCTGGCAGAGAAG 1845
|
|
|
QY 2226 AGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCAGTAGGAGGATGCACAGTTTGTCTATT 2285
|
|
|
Db 1846 AGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCAGTAGGAGGATGCACAGTTTGTCTATT 1905
|
|
|
QY 2286 TGCTTTAGACAGAGGAGTGTATTAACAAGCTTAACATTTGGTGCAAGATTTGCCTTTGA 2345
|
|
|
Db 1906 TGCTTTAGACAGAGGAGTGTATTAACAAGCTTAACATTTGGTGCAAGATTTGCCTTTGA 1965
|
|
|
QY 2346 A 2346
|
|
|
Db 1966 A 1966
|
|
|
```

Search completed: March 2, 2003, 13:25:59
Job time : 125 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2003, 14:26:30 : Search time 2167.5 Seconds

(without alignments)
17544.177 Million cell updates/sec

Title: US-09-723-722a-42

Perfect score: 2348

Sequence: 1 ccatcgccgcctcacagc.....caagattgcctctgaatt 2348

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: gb_gss:*
 - 18: em_gss_hum:*
 - 19: em_gss_inv:*
 - 20: em_gss_pln:*
 - 21: em_gss_vrt:*
 - 22: em_gss_fun:*
 - 23: em_gss_mam:*
 - 24: em_gss_mus:*
 - 25: em_gss_other:*
 - 26: em_gss_pro:*
 - 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1649	70.2	3859	11 AK014464	AK014464 Mus muscu
2	990	42.2	1114	9 AF150387	AF150387 AF150387
3	954.2	40.6	2503	11 AK014390	AK014390 Mus muscu
4	819.6	34.9	913	14 BQ949536	BQ949536 AGENCOURT
5	740.6	31.5	1059	13 BI084083	BI084083 602869445
6	728	31.0	747	14 BQ018588	BQ018588 UT-H-DHL

C	7	723.8	30.8	865	13	BI084878	BI084878 602869445
C	8	713.8	30.4	761	13	BM016288	BM016288 603642760
C	9	688.6	29.3	700	9	AL570757	AL570757 AL570757
C	10	688.6	29.3	904	10	BE378929	BE378929 601237528
C	11	681.4	29.0	718	12	BF057767	BF057767 7K53ell.x
C	12	664.4	28.3	945	14	BQ669630	BQ669630 AGENCOURT
C	13	657.8	28.0	684	12	BF338951	BF338951 602036021
C	14	647.4	27.6	730	13	BM048418	BM048418 603625683
C	15	634.2	27.0	840	12	BE885725	BE885725 601508937
C	16	631.4	26.9	789	12	BF204338	BF204338 601866411
C	17	624	26.6	635	10	BE296749	BE296749 601175134
C	18	598.6	25.5	684	10	AV725513	AV725513 AV725513
C	19	595	25.3	623	13	BI828125	BI828125 603075385
C	20	594.8	25.3	862	13	BI157614	BI157614 602922587
C	21	579	24.7	720	12	BG288435	BG288435 602383404
C	22	574.2	24.5	649	12	BE837592	BE837592 RC2-EN009
C	23	563.8	24.0	644	12	BE783981	BE783981 601472451
C	24	556.4	23.7	939	12	BF203806	BF203806 601868788
C	25	533	22.7	836	9	AL544727	AL544727 AL544727
C	26	524	22.3	525	9	AI005033	AI005033 ou91b12.x
C	27	514.4	21.9	541	10	AW770546	AW770546 hl86c06.x
C	28	511.4	21.8	513	14	BQ637035	BQ637035 hs04g12.y
C	29	508	21.6	864	12	BE872035	BE872035 601448124
C	30	476.2	20.3	605	12	BE793449	BE793449 601588510
C	31	475.2	20.2	514	12	BF439471	BF439471 nab65a03
C	32	472	20.1	487	13	BM313609	BM313609 ig72h03.y
C	33	464	19.8	464	9	AA701598	AA701598 zi34h07.s
C	34	464	19.8	464	9	AI127789	AI127789 qc32c02.x
C	35	453.2	19.3	458	9	AL700814	AL700814 DKFZp686H
C	36	451.2	19.2	542	10	AV667139	AV667139 AV667139
C	37	447.4	19.1	537	13	BG927600	BG927600 HNC43-1-E
C	38	445	19.0	461	9	AL700831	AL700831 DKFZp686I
C	39	445	19.0	493	12	BF443234	BF443234 260654 MA
C	40	444	18.9	444	9	AI095556	AI095556 qb19f04.x
C	41	443	18.9	443	9	AI127822	AI127822 qc36f02.x
C	42	426.4	18.2	428	9	AI094243	AI094243 qa43c11.s
C	43	425.6	18.1	1092	13	BM454007	BM454007 AGENCOURT
C	44	424.2	18.1	573	12	BE799078	BE799078 601588381
C	45	423.6	18.0	447	10	AW015855	AW015855 UT-H-BIO-

ALIGNMENTS

RESULT 1	AK014464	3859 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK014464	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:4122401C04;beta-site APP cleaving enzyme, full insert sequence.			
ACCESSION	AK014464				
VERSION	AK014464.1	GI:12852334			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) 16 days embryo head cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci,P. and Hayashizaki,Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				


```
QY 479 CAGCTTTGTGGAGATGGTGGACAACTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGA 538
DB 603 CAGCTTTGTGGAGATGGTGGACAACTGAGGGGAAAGTCGGGCCAGGGCTACTATGTGGA 662
QY 539 GATGACCGTGGGAGAGCCCGCCAGAGAGCTCAACATCTCTGGTGATACAGGCAGCAGTAA 598
DB 663 GATGACCGTGGGAGAGCCCGCCAGAGAGCTCAACATCTCTGGTGATACAGGCAGCAGTAA 722
QY 599 CTTTGCAGTGGGTGCTGCCCGCCCGCCAGAGAGCTCAACATCTCTGGTGATACAGGCAGCAGTGC 658
DB 723 CTTTGCAGTGGGTGCTGCCCGCCCGCCAGAGAGCTCAACATCTCTGGTGATACAGGCAGCAGTGC 782
QY 659 CAGCACAATACCGGAGAGCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGCAAGTGGGA 718
DB 783 CAGCACAATACGAGAGCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGCAAGTGGGA 842
QY 719 AGGGAGCTGGGCACGAGCTGTAGCATCCCGCCAGTCCCGCCAGAGCTCACTGTGGTGC 778
DB 843 GGGGGAACCTGGGACCGAGCTGTGTAGCATCCCGCCAGTCCCGCCAGAGCTCACTGTGGTGC 902
QY 779 CAACATTGTGCTGCCATCACTGAATCAGACAAAGTTCTTTCATCAACGGCTCCCAACTGGGAAG 838
DB 903 CAACATTGTGCTGCCATCACTGAATCAGACAAAGTTCTTTCATCAATGGTTCCAACTGGGAGG 962
QY 839 CATCTGGGGCTGGCCTATGCTGAGATGTGCCAGGCTGAGAGCTCCCTGGAGGCTTTCTT 898
DB 963 CATCTAGGGCTGGCCTATGCTGAGATGTGCCAGGCTGAGAGCTCTTTGGAGGCTTTCTT 1022
QY 899 TGACTCTCTGGTAAAGCAGACCCAGCTTCCCAACTCTTCTCCCTCGAGCTTTTGGTGC 958
DB 1023 TGACTCTCTGGTAAAGCAGACCCAGCTTCCCAACTCTTCTCCCTCGAGCTTTTGGGCGC 1082
QY 959 TGGCTTCCCGCTCAACAGCTCAAGTGTGCTGCTGCGAGGAGCATGATCATTTGG 1018
DB 1083 TGGCTTCCCGCTCAACAGCAGCAGCAGCTGGGCTCGGTGGGAGGAGCATGATCATTTGG 1142
QY 1019 AGGTATCGACCACTCGCTGTACAGAGCAGTCTCTGGTATACCCCATCCGCGGAGTG 1078
DB 1143 TGGTATCGACCACTCGCTGTACAGAGCAGTCTCTGGTATACCCCATCCGCGGAGTG 1202
QY 1079 GTATTATGAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGAGTCTGAAATGCACTG 1138
DB 1203 GTATTATGAGGTGATCATTTGTGGGTGGAGATCAATGGTCAAGATCTCAAGATGCACTG 1262
QY 1139 CAAGGAGTACAACATATGACAAGAGCATTTGGACAGTGGGACACCAACCTTCGTTGCC 1198
DB 1263 CAAGGAGTACAACATAGCAGAGAGCATTTGGACAGTGGGACACCAACCTTCGTTGCC 1322
QY 1199 CAAGAAAGTGTGTAAGCTGCACTCAAAATCCATCAAGGCGAGCTCTCCACGAGAGATT 1258
DB 1323 CAAGAAAGTATTTGAAGTGCCTCAAGTCCATCAAGGCGAGCTCTCCGACGAGAGATT 1382
QY 1259 CCCTGATGGTTTCTGGCTAGGAGAGCAGTGTGCTGTGCAAGGCGAGGACACCCCTTG 1318
DB 1383 CCGGATGGCTTTTGGCTAGGAGAGCAGTGTGCTGTGCAAGGCGAGGACACCCCTTG 1442
QY 1319 GAACATTTTCCAGTCACTCACTCACTTAATGGGTGAGTTTACCAACAGTCTTCCTCG 1378
DB 1443 GAACATTTTCCAGTCACTCACTTAATGGGTGAGTTTACCAACAGTCTTCCTCG 1502
QY 1379 CATCACCATCTTCCGAGCAGCAATACCTGGCGGAGTGGAAAGTGTGCCACGCTCCCAAGA 1438
DB 1503 CATCACCATCTTCCGAGCAGCAATACCTACGCGCGGTGGAGAGCTGCGCCACGCTCCCAAGA 1562
QY 1439 CGACTGTACAGTTTGGCCATCTACAGTCACTCCAGGCGAGCTTTATGGAGCTGTAT 1498
DB 1563 CGACTGTACAGTTTGGCCATCTACAGTCACTCCAGGCGAGCTTTATGGAGCGCTCAT 1622
QY 1499 CATGGAGGCTTCTAGCTTCTCTTGTGATCGGCGCGGAAAGCAATGGCTTTCTGTGCAG 1558
DB 1623 CATGGAGGCTTCTAGCTTCTCTTGTGATCGGCGCGGAAAGCAATGGCTTTCTGTGCAG 1682
QY 1559 CGCTTGCCATGTGCAGCATGAGTTTCAGGAGCGGAGCGGTGGAAAGGCCCTTTTGTACCTT 1618
```

```
DB 1683 CGCTTGCCATGTGCAGATGAGTTTCAGGACGGCGGAGTGAAGGTCCTTTGTACGGC 1742
QY 1619 GGACATGGAAGACTGTGGCTACAACTTCACAGACAGATGACTCAACCTCATGACCAT 1678
DB 1743 AGACATGGAAGACTGTGGCTACAACTTCACAGACAGATGAGTCAACATTTATGACCAT 1802
QY 1679 AGCTATGTCATGGCTGCCATCTCGGCCCTCTTTCATGCTGCCACTCTGCTCATGTGTG 1738
DB 1803 AGCTATGTCATGGCGGCATCTCGGCCCTCTTTCATGCTGCCACTCTGCTCATGTGTG 1862
QY 1739 TCAGTGGCGTGCCTCGCTGCTGCTGCCAGCAGCATGAGCTTTGCTGATGACATCTC 1798
DB 1863 TCAGTGGCGTGCCTCGCTGCTGCTGCCAGCAGCATGAGCTTTGCTGATGACATCTC 1922
QY 1799 CTTGCTGAAGTGAAGGAGCCCATGGCAGAGATGAGATTTCCCTGGACCCACACCTCCG 1858
DB 1923 CTTGCTGAAGTGAAGGAGCCCGTGGCAGATGATGAGAGCGCCCTGGACCATCTGGG 1982
QY 1859 TGGTTTCACCTTTGGTGCACA--AGTAGGAGACAGATGGCACCTGTGGCCAGACACCTCA 1916
DB 1983 TGGTTTCACCTTTGGTGCACATGAGTTGGAGCTATGATGGTACCTGTGGCCAGACACCTCA 2042
QY 1917 GGACCTTCCCGCCACCCCAAAATGCTGCTGCTGATGGAGAGAAAGGCTGGCAAGGT 1976
DB 2043 GGACCTTCCCGCCACCCCAAAATGCTGCTGCTGATGGAGAGAAAGGCTGGCAAGGT 2101
QY 1977 GGCCTTCCAGGAGCTGTACCTGTAGGAACAGAAAGAGAGAAAGAGCACTCTGCTGGC 2036
DB 2102 GGATTTACAGGCTTGGACCTGTAGGACACAGGAGAGGAGGAGGAGCGCTTTGGTGGC 2161
QY 2037 GGAATACTCTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTGTGAAACTTCAGC 2096
DB 2162 AGGAATATCTTTAGACACCAAACTTGAGT-TGGAATTTTCTGCTTGAAGCTTCAGC 2220
QY 2097 CTTGAGAACTTTGTCACCAATTTCCCTTTAAATTTCCCAACCCAAAGATATCTTTCTTTA 2156
DB 2221 CTTGAGCTCTGCGCCAGCGCTCTTTAGAGTCTCCAACTTAAAGTATTTTATGTCT-- 2277
QY 2157 GTTTCAGAACTACTGGCATCACAGCGAGTTACTTTGGCGTGTGCTGCTGGTACCCCTG 2216
DB 2278 CTTTCAGAACTACTGGCGTCTATCTCAGGCTACC-CGGCATGTGCTCCTGTGGTACCCCTG 2336
QY 2217 GCAGAGAAGACCAAGCTTTGTTCCCTGTGCGCAAGTCACTAGGAGGAGTGCACAG 2276
DB 2337 GCAGAGAAGGCGCAATC-TCATTCCTGTGCGCAAGTTCAGCAGAGA--AAGTGAAG 2393
QY 2277 TTTGCTATTTGCTTTTAGACAGGAGCTGTATAACAAAGCTTAACATTTGCTGCAAGATT 2336
DB 2394 TTTGCCAGTTGCTTTTAGTATAGGAGCTGCAGACTCAAGCT-ACACTGGTACAAAGACT 2452
QY 2337 GCCTCTGA 2345
DB 2453 GCCTCTGA 2461
```

RESULT 2

AF150387

LOCUS

DEFINITION

AF150387 Human mRNA from cd34+ stem cells

AF150387

ACCESSION

VERSION

AF150387.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1114)

AUTHORS

Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L.,

Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.

Human mRNA from cd34+ stem cells

Unpublished (1999)

AF150387 1114 bp mRNA linear EST 22-JUN-1999
AF150387 Human mRNA from cd34+ stem cells Homo sapiens cDNA clone
CBMAPH03, mRNA sequence.

AF150387.1 GI:5133823

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1114)

AUTHORS

Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L.,

Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.

Human mRNA from cd34+ stem cells

Unpublished (1999)

COMMENT

Contact: Ye M
Shanghai Institute of Hematology
Shanghai Second Medical University, Rui-Jin Hospital
197 Rui-Jin Road II, Shanghai, 200025, P. R. China
Email: zchen@stn.sh.cn

FEATURES
source

Location/Qualifiers
1. .1114
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CBMAPH03"
/clone_lib="Human mRNA from cd34+ stem cells"
/tissue_type="umbilical cord blood"
/cell_type="cd34+ stem cell"
/note="cloned by differential display method after
chemical induction of terminal differentiation of cell
line"

BASE COUNT 275 a 300 c 286 g 253 t

ORIGIN

Query Match 42.2%; Score 990; DB 9; Length 1114;
Best Local Similarity 97.3%; Pred. No. 4.8e-232;
Matches 1092; Conservative 0; Mismatches 20; Indels 10; Gaps 8;

QY 1017 GGAGGTATCGACCACTCGCTGTACACAGCAGCTCTCTGGTATACACCCATCCGGCGGGAG 1076
|||||
Db 1 GGAGGTATCGACCACTCGCTGTACACAGCAGCTCTCTGGTATACACCCATCCGGCGGGAG 60
|||||

QY 1077 TGGTATTATGAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGAC 1136
|||||
Db 61 TGGTATTATGAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGAC 120
|||||

QY 1137 TGAAGAGTACAACTATGACAAGAGCATTTGGACAGTGGACACCAACCTTCGTTTG 1196
|||||
Db 121 TGAAGAGTACAACTATGACAAGAGCATTTGGACAGTGGACACCAACCTTCGTTTG 180
|||||

QY 1197 CCCAAGAAAGTGTGAAGCTGCAGTCAATCCATCAAGCAGCCTCTCCACGGAGAAG 1256
|||||
Db 181 CCCAAGAAAGTGTGAAGCTGCAGTCAATCCATCAAGCAGCCTCTCCACGGAGAAG 240
|||||

QY 1257 TTCCCTGATGTTTCTGGCTAGGAGAGCAGCTGGTGTGGCAAGCAGGACCAACCCCT 1316
|||||
Db 241 TTCCCTGATGTTTCTGGCTAGGAGAGCAGCTGGTGTGGCAAGCAGGACCAACCCCT 300
|||||

QY 1317 TGAACATTTTCCAGTCATCTCACTTACCTAATGGGTGAGTTACCAACAGCTCTTC 1376
|||||
Db 301 TGAACATTTTCCAGTCATCTCACTTACCTAATGGGTGAGTTACCAACAGCTCTTC 360
|||||

QY 1377 CGCATCAACCTCTCCGACCAATACCTCGCGCCAGTGGAGATGTGGCCACGTCCTCAA 1436
|||||
Db 361 CGCATCAACCTCTCCGACCAATACCTCGCGCCAGTGGAGATGTGGCCACGTCCTCAA 419
|||||

QY 1437 GACGACTGTTTACAAAGTTTGGCATCTCACAGTCATCCACGGGCACTGTTATGGAGCTGTT 1496
|||||
Db 420 GACGACTGTTTAC - AGTTTGGCATCTCACAGTCATCCACGGGCACTGTTATGGAGCTGTT 478
|||||

QY 1497 ATCATGGAGGCTTCTA - CGTTGCTTTGATCGGGCCCAACAGATTTGGCTTTGCTGT 1555
|||||
Db 479 ATCATGGAGGCTTCTA - CGTTGCTTTGATCGGGCCCAACAGATTTGGCTTTGCTGT 538
|||||

QY 1556 CAGCGCTTGCATCTGCACGATGAGTTTCAGGACGGCAGCGGTGGAGGCGCTTTTGTTCAC 1615
|||||
Db 539 CAGCGCTTGCATCTGCACGATGAGTTTCAGGACGGCAGCGGTGGAGGCGCTTTTGTTCAC 597
|||||

QY 1616 CTTGGACATGGAAGACTGTGGCTACAACATTCACAGACAGATGAGTCAACCTTCATGAC 1675
|||||
Db 598 CTTGGACATGGAAGACTGTGGCTACAACATTCACAGACAGATGAGTCAACCTTCATGAC 657
|||||

QY 1676 CATAGCCTATGTCATGGTGCATCTGGCCCTCTTTCATGCTGCCACTTGCCTCATGGT 1735
|||||
Db 658 CATAGCCTATGTCATGGTGCATCTGGG - CCTCTTCATGCTGCCACTTGCCTCATGGT 716
|||||

QY 1736 GTCTCAG - TGGCGTGCCTCGCTCGCTCGCTCGCCAGCAGCATGATGATTTGCTGCATGACA 1794
|||||

Db 717 GTGTGAGATGGCGTGCCTTCGCTGCCTTCGCGCAGCAGCATGATGACTATCTGTATGACA 776
|||||

QY 1795 TCTCCCTGCTCAAGTGAGGAGGCCATGGGCACAAGATAGAGATTCCCTCTGGACACACC 1854
|||||

Db 777 TCTCCCTGCTCAAGTGAGGAGGCCATGGGCACAAGATAGAGATTCCCTCTGGACACACC 836
|||||

QY 1855 TCCGTGGTTCACTTTTGGTTCACAAAGTAGGACACAGATGCGACCTGTGGCCAGACACCT 1914
|||||

Db 837 TACGTGGTCACTTTTGGCAC - - - AGTAGGAGACACAGATGCGACCTGTGGCCAGACACCT 893
|||||

QY 1915 CAGGACCTCTCCACACCCACCAATGCCTCTGCCTTGTATGAGAGAAGAAAGCTGGCAAG 1974
|||||

Db 894 CAGGACCTCTCCACACCCACCAATGCCTCTGCCTTGTATGAGAGAAGAAAGCTGGCAAG 953
|||||

QY 1975 GTGGGTTCCAGGGAGCTGTACCTGTAGGAAACACAAAGAGAAGAAAGCACTCTGCTG 2034
|||||

Db 954 GTGGGTTCCAGGGAGCTG - ACCTGTAGGAAACACAAAGAGAAGAAAGCACTCTGCTG 1012
|||||

QY 2035 GCGGAATACCTTTGGTTCACCTCAATTTAAGTCGGGAATTTCTGCTGCTTTGAAACTTCA 2094
|||||

Db 1013 GCGCAATACCTTTGGTTCACCTCAATTTAAGTCGGGAATTTCTGCTGCTTTGAAACTTCA 1072
|||||

QY 2095 GCGCTGAACCTTTGTCCACCATTCCTTTAAATTTCTCAACCC 2136
|||||

Db 1073 ACCTGAACCTTTGTCAACCATTCCTTTAAATTTATACAACCC 1114
|||||

RESULT 3
AK014390

LOCUS
DEFINITION
2503 bp mRNA linear HTC 19-JAN-2002
Mus musculus adult male brain cDNA, RIKEN full-length enriched
library, clone:3526402A15:beta-site APP cleaving enzyme, full
insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AK014390
AK014390.1 GI:12852207
HTC; CAP trapper
Mus musculus (strain: C57BL/6J) adult male brain cDNA to mRNA,
clone_11b:RIKEN full-length enriched mouse cDNA library
clone:3526402A15.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Tanaka, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, T.,


```

QY 1729 TCATGTGTGTCTAGTGGCGCTCCCTCCCTGCTGCTGCGCAGCAGCATGATGACTTTGCTG 1788
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 721 TCATGTGTGTCTAGTGGCGCTCCCTCCCTGCTGCTGCGCAGCAGCATGATGACTTTGCTG 780

QY 1789 ATGACATCTCCCTGCTGAAGTGTAGGAGGCCATGGGCAGAAATAGAGATTCCTCCCTGGAC 1848
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 781 ATGACATCTCCCTGCTCAAGTAAGGAGGCCCTGGGCAGATGATGGAGCGCCCTGGAC 840

QY 1849 CACACCTCCCTGGTTCACCTTTGGTTCACA - AGTAGGAGACACAGATGCGCACTTGCGCCA 1906
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 841 CACATCTGGTGGTTCCTTTGGTTCACATGAGTTGGAGCTATGGATGGTACCTGTGGCCA 900

QY 1907 GAGCACCCTAGGACCCCTCCCAACCCACCAAAATGCCTCTCCCTTGTATGGAGAAAGAAAGG 1966
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 901 GAGCACCCTAGGACCCCTCCCAACCCACCAAAATGCCTCTCCCTTGTATGGAGAAAGAAAG 959

QY 1967 CTGGCAAGTGGTTCAGGGACTGTACCTGTAGGAACAGAAAGAAAGAAAGCA 2026
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 960 CAGGCAAGCTGGATTCAGGGCTTTCACCTGTAGGACACAGGAGAGGAAAGCAAGCAGCG 1019

QY 2027 CTCTGCTGGCGGGAATACTCTTGGTTCACCTCAAAATTAAGTCGGGAAATCTGCTGCTTG 2086
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1020 TTCTGCTGGCAGGAATATCTCTAGACACCAAACTTGAGT - TGGAAATTTTGTGCTTG 1078

QY 2087 AAACCTCAGCCCTGAACCTTTTGTCCACCATTCCTTTAAATTTCCAAACCCAAAGTATCT 2146
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1079 AAGCTTCAGCCCTGACCTCTGCGCCAGCA - TCCTTTAGAGTCTCCAACCTTAAAGTATCT 1137

QY 2147 TCCTTTCTTAGTTTCAGAAGTACTGGCATCACACGAGGTACCTTGGCGTGTGCTCCTG 2206
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1138 TTATGTC - -CTTCCAGAAGTACTGGCGTCATACCTAGCTAGC - CGGCGATGTGCTCCTG 1193

QY 2207 TGTACCTGGCAGAGAGACCAAGCTGTGTTCCCTGCTGGCCAAAGTCAGTAGGAGA 2266
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1194 TGTACCTGGCAGAGAGAGGCGCAATC - TCATTTCCCTGCTGGCCAAAGTCAGCAGAGA 1252

QY 2267 GGATGCACAGTTGCTATTTGCTTTAGAGACAGGACTGTATAAACAAGCTTAACATTTGG 2326
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1253 -AAGTGAAGTTTCCAGTTGCTTTAGTGATAGGAGCTGCAGACTCAAGCCT - ACACCTG 1309

QY 2327 TGCAAGATTTGCCTCTTGA 2345
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1310 TACAAGACTGCGGTCTTGA 1328

RESULT 4
BQ949536
LOCUS
DEFINITION BQ949536 913 bp mRNA linear EST 21-AUG-2002
5', mRNA sequence.
AGENCOURT_8733256 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6454803
ACCESSION BQ949536
VERSION BQ949536.1 GI:22365014
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 913)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2632 row: h column: 04
High quality sequence stop: 613.

```

```

FEATURES
Source
Location/Qualifiers
1..913
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pORF7; Site: 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 207 a 259 c 232 g 206 t
ORIGIN
Query Match 34.9%; Score 819.6; DB 14; Length 913;
Best Local Similarity 96.8%; Pred. No. 3e-190;
Matches 864; Conservative 0; Mismatches 26; Indels 3; Gaps 3;
QY 1209 TTTGAAGCTGAGTCAAAATCCATCAAGGAGGCTCTCCACGGAGAGATTCCCTGATGCT 1268
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 TTTGAAGCTGAGTCAAAATCCATCAAGGAGGCTCTCCACGGAGAGATTCCCTGATGCT 60

QY 1269 TTTCTGGCTAGGAGAGCAGCTGCTGCTGCGCAAGCAGGACACCCCTTGGAAACATTTTC 1328
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 TTTCTGGCTAGGAGAGCAGCTGCTGCTGCGCAAGCAGGACACCCCTTGGAAACATTTTC 120

QY 1329 CCAGTCATCTCAGTCTACCTAATGGGTGAGTTTACCAACAGCTCTTCCGCATCACCATC 1388
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 CCAGTCATCTCAGTCTACCTAATGGGTGAGTTTACCAACAGCTCTTCCGCATCACCATC 180

QY 1389 CTTCCGCGAGCAATACCTGCGGCGCAGTGAAGATGTGGGCAGCTGCCAAGAGACTGTGTAC 1448
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 CTTCCGCGAGCAATACCTGCGGCGCAGTGAAGATGTGGGCAGCTGCCAAGAGACTGTGTAC 240

QY 1449 AAGTTTCCCATCTCACAGTCATCCACGGCAGCTGTTTATGGAGCTGTTATCATCGAGGCG 1508
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 AAGTTTCCCATCTCACAGTCATCCACGGCAGCTGTTTATGGAGCTGTTATCATCGAGGCG 300

QY 1509 TTCTACGCTTGTCTTGTATCGGCGCCGAAACGAATTTGGCTTGTCTGACGCTTGGCAT 1568
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 TTCTACGCTTGTCTTGTATCGGCGCCGAAACGAATTTGGCTTGTCTGACGCTTGGCAT 360

QY 1569 GTCCACGATGAGTTCAGGACGCGCAGCGGTGGAAGGCCCTTTTCTCAGCTTGGACATGGAA 1628
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 361 GTCCACGATGAGTTCAGGACGCGCAGCGGTGGAAGGCCCTTTTCTCAGCTTGGACATGGAA 420

QY 1629 GACTGTGGCTACACATTTCCACAGACAGATGATCAACCTCATGACCATAGCCTATCTC 1688
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 421 GACTGTGGCTACACATTTCCACAGACAGATGATCAACCTCATGACCATAGCCTATCTC 480

QY 1689 ATGGCTGCCATCTCGCCCTCTTTCATGCTGCCACTCTGCTCATGTTGTGTGCTGAGGCG 1748
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 481 ATGGCTGCCATCTCGCCCTCTTTCATGCTGCCACTCTGCTCATGTTGTGTGCTGAGGCG 540

QY 1749 TGCTTCGGCTGCTCGGCGCAGCAGCATGATGATTTTGTGTGATGACATCTCCTCTCAAG 1808
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 541 TGCTTCGGCTGCTCGGCGCAGCAGCATGATGATTTTGTGTGATGACATCTCCTCTCAAG 600

QY 1809 TGAGGAGGCCCATCGGCGAGAGATAGAGATTCCCTTGGACACACCTCCCTGGTTCATT 1868
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 601 TGAGGAGGCCCATCGGCGAGAGATAGAGATTCCCTTGGACACACCTCCCTGGTTCATT 660

QY 1869 TGCTCACAAAGTAGGAGACACAGATGGCAGCTGTGGCCAGAGACCTCAGACCTCCGCCA 1928
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 661 TGCTCACAAAGTAGGAGACACAGATGGCAGCTGTGGCCAGAGACCTCAGACCTCCGCCA 720

QY 1929 CCCACCAAAATGCTCTGCTTTGATGGAGAGAGGAAAAGGTGGCAAGGTGGTTCCTCA -GGG 1987
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```


cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 The following repetitive elements were found in this cDNA
 sequence: 1-22, >AT-rich#Low_complexity
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..747
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5823683"
 /clone_lib="NCI CGAP DHL"
 /tissue_type="Metastatic Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a
 modified polylinker; Site_1: Ecor I; Site_2: Not I;
 NCI CGAP DHL is a normalized cDNA library containing the
 following tissue(s): VS-8 Cell line from Metastatic
 Chondrosarcoma in Lung. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an Ecor I adaptor, digested
 with Not I, and cloned directionally into p773-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is AGATCATTCG.
 TAG_LIB=UI-H-DHL
 TAG_TISSUE=Lung
 TAG_SEQ=AGATCATTCG"

BASE COUNT 176 a 178 c 192 g 200 t 1 others
 ORIGIN

Query Match 31.0%; Score 728; DB 14; Length 747;
 Best Local Similarity 99.9%; Pred. No. 8.8e-168;
 Matches 728; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1620 GACATGGAAGACTGTGCTACAACTCCAGACAGATGATCAACCCCTCATGACCATA 1679
 Db 747 GACATGGAAGACTGTGCTACAACTCCAGACAGATGATCAACCCCTCATGACCATA 688
 Qy 1680 GCCTATGTCATGCTGCCTCCTTCATGCTGCCACTCTGCCTCATGGTGTGT 1739
 Db 687 GCNTATGTCATGCTGCCTCCTTCATGCTGCCACTCTGCCTCATGGTGTGT 628
 Qy 1740 CAGTGGCGTCCCTCGCTCGCCGACGACGATGATGATTTGCTGATGACATCTCC 1799
 Db 627 CAGTGGCGTCCCTCGCTCGCCGACGACGATGATGATTTGCTGATGACATCTCC 568
 Qy 1800 CTGCTGAAGTGAAGAGCCCATGGGACAGAGATAGAGATTCCTCCGACACACCTCCGT 1859
 Db 567 CTGCTGAAGTGAAGAGCCCATGGGACAGAGATAGAGATTCCTCCGACACACCTCCGT 508
 Qy 1860 GCTTCACTTTGTCACAAGTAGGAGACACAGATGACCTGTGGCCAGACACCTCAGGA 1919
 Db 507 GGTTCACCTTTGGTCACAAGTAGGAGACACAGATGACCTGTGGCCAGACACCTCAGGA 448
 Qy 1920 CCCTCCACCCACCAATGCTCTGCTTTCATGGAGAGGAAAGGCTGGCAAGTGGG 1979
 Db 447 CCCTCCACCCACCAATGCTCTGCTTTCATGGAGAGGAAAGGCTGGCAAGTGGG 388
 Qy 1980 TTCCAGGGACTGACCTGTAGGAACAGAGAAAGAGAAAGCACTCTGCTGGCGGG 2039
 Db 387 TTCCAGGGACTGACCTGTAGGAACAGAGAAAGAGAAAGCACTCTGCTGGCGGG 328
 Qy 2040 ATACTCTTGGTCACTCAATTTAAGTCGGGAATTCGTGCTTGAACCTTCAGCCCT 2099
 Db 327 AATACTCTTGGTCACTCAATTTAAGTCGGGAATTCGTGCTTGAACCTTCAGCCCT 268

Qy 2100 GAACCTTTGTCCACCATTTCCTTTAAATTTCTCCAAACCAAGATATCTCTTTCTTACTTT 2159
 Db 267 GAACCTTTGTCCACCATTTCCTTTAAATTTCTCCAAACCAAGATATCTCTTTCTTACTTT 208
 Qy 2160 TCAGAACTACTGGCATCATCACAGCAGGTACCTTGGCGTGTGTCCTGTGTGACCTGCGCA 2219
 Db 207 TCAGAACTACTGGCATCATCACAGCAGGTACCTTGGCGTGTGTCCTGTGTGACCTGCGCA 148
 Qy 2220 GAGAAGAGACCAAGCTTGTTCCTTCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTT 2279
 Db 147 GAGAAGAGACCAAGCTTGTTCCTTCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTT 88
 Qy 2280 GCTATTTCCTTTAGACAGCAGGACTGTATAAACAAGCCCTAAACATTTGCTGCAAAAGATGCC 2339
 Db 87 GCTATTTCCTTTAGACAGCAGGACTGTATAAACAAGCCCTAAACATTTGCTGCAAAAGATGCC 28
 Qy 2340 TCTTGAATT 2348
 Db 27 TCTTGAATT 19

RESULT 7
 BI084878/c
 LOCUS
 DEFINITION 602869445T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013994 3',
 mRNA sequence.
 ACCESSION BI084878
 VERSION BI084878.1 GI:14503208
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-femail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCMI820 row: j column: 11
 High quality sequence start: 18
 High quality sequence stop: 821.
 Location/Qualifiers
 1..865
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5013994"
 /clone_lib="NIH_MGC_102"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies). Note: this is a NIH-MGC Library."

FEATURES

Source
 1..865
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5013994"
 /clone_lib="NIH_MGC_102"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 207 a 224 c 228 g 205 t
 ORIGIN
 Query Match 30.8%; Score 723.8; DB 13; Length 865;
 Best Local Similarity 95.3%; Pred. No. 9.7e-167;
 Matches 810; Conservative 0; Mismatches 33; Indels 7; Gaps 6;


```

Db 541 CTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTGAAGTGAGGAGGCCCA 600
QY 1821 TGGCAGAGATAGAGATTCCCTTGACACACACCTCGGTGTTTACATTTGCTCAAGTA 1880
|||||
Db 601 TGGCAGAGATAGAGATTCCCTTGACACACACCTCGGTGTTTACATTTGCTCAAGTA 660
|||||
QY 1881 GGAGACACAGATGGCAGCTGTGGCCAGAGACACCTCAGGACCCTCCACCCACCAAAATGC 1940
|||||
Db 661 GGAGACACAGATGGCAGCTGTGGCCAGAGACACCTCAGGACCCTCCCA-CCACAAATGC 719
|||||
QY 1941 CTCTGCCCTTGA 1951
|||||
Db 720 CTCTGCCCTTGA 730
|||||

RESULT 9
AL570757/c
LOCUS
DEFINITION AL570757 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI022YF22 3
prime, mRNA sequence.
ACCESSION AL570757
VERSION AL570757.1 GI:12927378
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Pollayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI022YF22"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 169 a 173 c 191 g 161 t 6 others
ORIGIN

Query Match 29.3%; Score 688.6; DB 9; Length 700;
Best Local Similarity 99.3%; Pred. No. 4e-158;
Matches 688; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1555 TCAGCGCTTGCATGTGCACGATGATTCAGGACGCGCGGTGGAAGGCCCTTTGTGCA 1614
|||||
Db 693 TCAGCGCTTGCATGTGCACGATGATTCAGGACGCGCGGTGGAAGGCCCTTTGTGCA 634
|||||
QY 1615 CTTTGACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGA 1674
|||||
Db 633 CTTTGACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGA 574
|||||
QY 1675 CCATAGCCTATGTGATGGCTGCCATCTTGGCCGCTCTTTCATGTGCGCACTCTGCCCTCATGG 1734
|||||
Db 573 CCATAGCCTATGTGATGGCTGCCATCTTGGCCGCTCTTTCATGTGCGCACTCTGCCCTCATGG 514
|||||
QY 1735 TGTGTAGTGGCGCTGCTCCGCTCGCCGCGCAGCAGCATGATGACTTGTCTGATGACA 1794
|||||

```

```

Db 513 TGTGTAGTGGCGCTGCCCTCCGCTGCTGCGCCACAGCATGATGACTTGTGCTGATGACA 454
QY 1795 TCTCCCTCTGAAGTTCAGAGGCCCATGGCAGAGATAGAGATTCCCTTGACACACACC 1854
|||||
Db 453 TCTCCCTCTGAAGTTCAGAGGCCCATGGCAGAGATAGAGATTCCCTTGACACACACC 394
|||||
QY 1855 TCCGTGGTTCACATTTGGTTCACAAAGTAGGAGACACAGATGGCACCTTGTGGCCAGAGACCT 1914
|||||
Db 393 TCCGTGGTTCACATTTGGTTCACAAAGTAGGAGACACAGATGGCACCTTGTGGCCAGAGACCT 334
|||||
QY 1915 CAGGACCTCCGCCACCCACCAATGGCTCTGCTTGTAGGAGAGAAAGGCTGGCAAG 1974
|||||
Db 333 CAGGACCTCCGCCACCCACCAATGGCTCTGCTTGTAGGAGAGAAAGGCTGGCAAG 274
|||||
QY 1975 GTGGTTCCAGGAGCTGTACCTGTAGGAAACAGAAAGAGAAAGAGCACTCTGCTG 2034
|||||
Db 273 GTGGTTCCAGGAGCTGTACCTGTAGGAAACAGAAAGAGAAAGAGCACTCTGCTG 214
|||||
QY 2035 GCGGGAATACTCTTGTGTACCTCAAAATTTAAGTCGGGAAATTTCTGCTGTTGAAACTTCA 2094
|||||
Db 213 GCGGGAATACTCTTGTGTACCTCAAAATTTAAGTCGGGAAATTTCTGCTGTTGAAACTTCA 154
|||||
QY 2095 GCCGTGAACCTTTGTGCCACCATTCCTTTAAATTTCTCAACCCAAAGATTTCTTTTCT 2154
|||||
Db 153 GCCGTGAACCTTTGTGCCACCATTCCTTTAAATTTCTCAACCCAAAGATTTCTTTTCT 94
|||||
QY 2155 TAGTTTCAGAGTACTTGGCATCACACGAGTTACCTTGGCGTGTGTCCTGTGTGTAACC 2214
|||||
Db 93 TAGTTTCAGAGTACTTGGCATCACACGAGTTACCTTGGCGTGTGTCCTGTGTGTAACC 34
|||||
QY 2215 TGGCAGAGAGAGACCAAGCTTTGTTCCCTGCT 2247
|||||
Db 33 TGGCAGAGAGAGAGACCAAGCTTTGTTCCCTGCT 1

RESULT 10
BE378929
LOCUS
DEFINITION 601237528F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3609631 5',
mRNA sequence.
ACCESSION BE378929
VERSION BE378929.1 GI:9324294
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM264 row: g column: 08
High quality sequence stop: 735.
Location/Qualifiers
1..904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3609631"
/clone_lib="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="pH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'

```


SOURCE: human.

FEATURES
source
High quality sequence stop: 652.
Location/Qualifiers
1. 684
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_Brn64"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/note="organ: brain; Vector: PCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 161 a 189 c 172 g 162 t
ORIGIN

Query Match 28.0%; Score 657.8; DB 12; Length 684;
Best Local Similarity 99.4%; Pred. No. 1.5e-150;
Matches 681; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1534 GAAAGCAATTGGCTTGTGTCAGCGCTTGCATGTGCACGATGAGTTTCAGGACGGCAG 1593
Db 1 GAAAGCAATTGGCTTGTGTCAGCGCTTGCATGTGCACGATGAGTTTCAGGACGGCAG 60
QY 1594 CGGTGAAGCGCCCTTTGTACCTTGGACATGGAAGACTGTGGCTACAACTTCCACAGA 1653
Db 61 CGGTGAAGCGCCCTTTGTACCTTGGACATGGAAGACTGTGGCTACAACTTCCACAGA 120
QY 1654 CAGATGAGTCAACCTTCATGACATAGCCCTATGTATGCTGCGCATCTGCGCCCTTTCA 1713
Db 121 CAGATGAGTCAACCTTCATGACATAGCCCTATGTATGCTGCGCATCTGCGCCCTTTCA 180
QY 1714 TGCTGCCACTCTGCCATCATGTTGTACGTGGCGTCCCTCGCTCGCTCGCCACGAGC 1773
Db 181 TGCTGCCACTCTGCCATCATGTTGTACGTGGCGTCCCTCGCTCGCTCGCCACGAGC 240
QY 1774 ATGATGACTTTGCTGATGACATCTCCCTGCTGCTGAGTGGAGGAGGCCCATGGCAGAGA 1833
Db 241 ATGATGACTTTGCTGATGACATCTCCCTGCTGCTGAGTGGAGGAGGCCCATGGCAGAGA 300
QY 1834 GAGATTCCCTGGACACACCTCCCGTGGTTCACCTTGGTCAACAAGTAGGAGACACAGATG 1893
Db 301 GAGATTCCCTGGACACACCTCCCGTGGTTCACCTTGGTCAACAAGTAGGAGACACAGATG 360
QY 1894 GCACCTGTGGCCAGACACCTCAGGACCTCCACCACCCACCAATCCCTCTGCTTTGATG 1953
Db 361 GCACCTGTGGCCAGACACCTCAGGACCTCCACCACCCACCAATCCCTCTGCTTTGATG 420
QY 1954 GAGAAGAAAGGCTGGCAAGGTGGTTCAGGGACTGTACCTGTAGGAAACAGAAAAGA 2013
Db 421 GAGAAGAAAGGCTGGCAAGGTGGTTCAGGGACTGTACCTGTAGGAAACAGAAAAGA 479
QY 2014 GAAGAAGAAAGCTGCTGGCGGGAATACCTTGGTCAACCTCAATTTAAGTCCGGAA 2073
Db 480 GAAGAAGAAAGCTGCTGGCGGGAATACCTTGGTCAACCTCAATTTAAGTCCGGAA 539
QY 2074 ATTCTGCTGCTGAACACTTCAGCCCTGGAACCTTTGTCACCATTCCTTTAAATTCACCA 2133
Db 540 ATTCTGCTGCTGAACACTTCAGCCCTGGAACCTTTGTCACCATTCCTTTAAATTCACCA 599
QY 2134 CCC-AAAGTATTCTTCTTTCTTTAGTTTCAGAAAGTACTGGCATCACACGAGTTTACCTT 2192
Db 600 CCCAAAGTATTCTTCTTTCTTTAGTTTCAGAAAGTACTGGCATCACACGAGTTTACCTT 659
QY 2193 GCGGTGTCCCTGTGGTACCCCTGG 2217
Db 660 GCGGTGTCCCTGTGGTACCCCTGG 684

RESULT 14
BM048418 730 bp mRNA linear EST 07-NOV-2001
LOCUS 603625683F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452052 5',
DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BM048418
BM048418.1 GI:16777685
EST
human.
Homo sapiens

mrna sequence.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 730)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue procurement: DCTD/Drp
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1943 row: f column: 21
High quality sequence stop: 696.
Location/Qualifiers
1. 730
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5452052"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site.1: XhoI;
Site.2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 163 a 210 c 206 g 151 t
ORIGIN

Query Match 27.6%; Score 647.4; DB 13;
Best Local Similarity 97.0%; Pred. No. 5.2e-148;
Matches 713; Conservative 0; Mismatches 16; Indels 6; Gaps 5;

QY 539 GATGACCGTGGGCGAGCCCGCCGAGCGCTCAACATCTCGTGGATACAGCGCAGCTAA 598
Db 2 GATGACCGTGGGCGAG-CCCGCCGAGCGCTCAACATCTCGTGGATACAGCGCAGCTAA 60
QY 599 CTTTTCAGTGGGTGCTGCCCGCCCGCCCGCTTCCCTGCATCGCTACTACCGAGGCGCTGC 658
Db 61 CTTTTCAGTGGGTGCTGCCCGCCCGCCCGCTTCCCTGCATCGCTACTACCGAGGCGCTGC 120
QY 659 CAGCACATACCGGGACCTCCGGAAGGTGTGTATGTGCCCTACACCCAGGGCAGGTGGGA 718
Db 121 CAGCACATACCGGGGA-CTCCGGAAGGTGTGTATGTGCCCTACACCCA-GGCAAGTGGGA 178
QY 719 AGGGGAGCTGGGCGACCGACCTGTGAAGCATCCCGCCCGCCCGCCCGCTACTGTGCGTGC 778
Db 179 AGGGGAGCTGGGCGACCGACCTGTGAAGCATCCCGCCCGCTACTGTGCGTGC 237
QY 779 CAACATTGCTGCCATCCTACTGAATCAGACAAGTTCTTCATCAACGGCTCCCACTGGGAAG 838
Db 238 CAACATTGCTGCCATCCTACTGAATCAGACAAGTTCTTCATCAACGGCTCCCACTGGGAAG 297
QY 839 CATCTCTGGGCTGGCCCTATGCTGAGATTGCCAGGCTGAGACTCCCTGGAGGCTTTCTT 898
Db 298 CATCTCTGGGCTGGCCCTATGCTGAGATTGCCAGGCTGAGACTCCCTGGAGGCTTTCTT 357
QY 899 TGACTCTCTGGTAAAGCAGACACCGCTTCCCAACCTCTTCTCCCTCGAGCTTTGTGTGTC 958
Db 358 TGACTCTCTGGTAAAGCAGACACCGCTTCCCAACCTCTTCTCCCTCGAGCTTTGTGTGTC 417

Qy	959	TGGCTTCCCCCTCAACCCAGTCTGAAGTGTCTGGCCCTCTGTGCGAGGAGCATGATCATGG	1018
Db	418	TGGCTTCCCCCTCAACCCAGTCTGAAGTGTCTGGCCCTCTGTGCGAGGAGCATGATCATGG	477
Qy	1019	AGGTATCGACCACTCGGTGTACACAGCAGTCTCGTGTATACACCCATCCGCGGGAGTG	1078
Db	478	AGGTATCGACCACTCGGTGTACACAGCAGTCTCGTGTATACACCCATCCGCGGGAGTG	537
Qy	1079	GTATTATGAGGTGATCATTTGTGCGGTGGAGATCAATGCACAGGATCTGAAAATGGACTG	1138
Db	538	GTATTATGAGGTGATCATTTGTGCGGTGGAGATCAATGCACAGGATCTGACACTGGACTG	597
Qy	1139	CAAGGAGTCAACTATGACAAGACATTTGTGGACAGTGGCACACCAACCTTCTGTTTGGC	1198
Db	598	CAAGGAGTCAACTATGACAAGACATTTGTGGACAGTGGCACAC--AACCTCGGATGCC	655
Qy	1199	CAAGAAAGTCTTTGAAGCTCGACTCAATTCATCAAGCAGCCTCTCTCACGGAGAGTT	1258
Db	656	CAAGAAAGTCTCGGAAGCTCGACTCAATTCATCAAGCAGCCTCTCTCACGGAGAGGT	715
Qy	1259	CCCTGATGTTTCTG 1273	
Db	716	CCCTGATGTTTCTG 730	
RESULT 15			
BE885725			
LOCUS			
DEFINITION	BE885725	840 bp mRNA linear EST 20-OCT-2000	
ACCESSION	601508937F1 NIH_MGC_71	Homo sapiens cDNA clone IMAGE:3910376 5',	
VERSION	BE885725	mRNA sequence.	
KEYWORDS	BE885725.1	GI:10339345	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
COMMENT	1 (bases 1 to 840)		
	NIH-MGC http://mgi.nci.nih.gov/.		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs@mail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone Distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM9725 row: j column: 09		
	High quality sequence stop: 628.		
FEATURES	Location/Qualifiers		
source	1..840		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3910376"		
	/clone_lib="NIH_MGC_71"		
	/tissue_type="lelomyosarcoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: uterus; Vector: pCMV-SPORT6; Site:1: NotI;		
	Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.		
	Average insert size 2.1 kb.		
BASE COUNT	167 a 216 c 248 g 209 t		
ORIGIN			
Query Match	27.0%;	Score 634.2;	DB 12; Length 840;
Best Local Similarity	94.4%;	Pred. No. 9.3e-145;	
Matches 590; Conservative	0; Mismatches 38;	Indels 3; Gaps 3;	
Qy	944	GCAGCTTTGTGGTGGCTTCCCTCAACCAAGTCTGAAGTGGCTGCGCTCTGCGAGG	1003

1	2041.4	86.9	2070	4	US-09-548-372D-3	Sequence 3, Appli
2	2041.4	86.9	2070	4	US-09-548-367D-3	Sequence 3, Appli
3	2024	86.2	2541	4	US-09-009-191-1	Sequence 1, Appli
4	1881	80.1	1977	4	US-09-548-372D-5	Sequence 5, Appli
5	1881	80.1	1977	4	US-09-548-367D-5	Sequence 5, Appli
6	1835.4	78.2	2370	4	US-09-009-191-3	Sequence 3, Appli
7	1550	66.0	2043	4	US-09-548-372D-7	Sequence 7, Appli
8	1550	66.0	2043	4	US-09-548-367D-7	Sequence 7, Appli
9	1355.8	57.7	1362	4	US-09-548-372D-29	Sequence 29, Appl
10	1355.8	57.7	1362	4	US-09-548-367D-29	Sequence 29, Appl
11	1355.8	57.7	1380	4	US-09-548-372D-31	Sequence 31, Appl
12	1355.8	57.7	1380	4	US-09-548-367D-31	Sequence 31, Appl
13	1298	55.3	1506	4	US-09-713-158-1	Sequence 1, Appli
14	1293.8	55.1	1341	4	US-09-548-372D-21	Sequence 21, Appl
15	1293.8	55.1	1341	4	US-09-548-367D-21	Sequence 21, Appl
16	1293.8	55.1	1380	4	US-09-548-372D-23	Sequence 23, Appl
17	1293.8	55.1	1380	4	US-09-548-367D-23	Sequence 23, Appl
18	1270.4	54.1	1302	4	US-09-548-372D-25	Sequence 25, Appl
19	1270.4	54.1	1302	4	US-09-548-367D-25	Sequence 25, Appl
20	1197.4	51.0	1287	4	US-09-548-372D-50	Sequence 50, Appl
21	1197.4	51.0	1287	4	US-09-548-367D-50	Sequence 50, Appl
22	1197.4	51.0	1305	4	US-09-548-372D-52	Sequence 52, Appl
23	1197.4	51.0	1305	4	US-09-548-367D-52	Sequence 52, Appl
24	1180	50.3	1278	4	US-09-548-372D-27	Sequence 27, Appl
25	1180	50.3	1278	4	US-09-548-367D-27	Sequence 27, Appl
26	421.2	17.9	511	4	US-09-280-116-30	Sequence 30, Appl
27	397.4	16.9	1804	4	US-09-548-372D-1	Sequence 1, Appli

Qy 606 GTGGTGTGCCCCCACCACCCCTTCTCGATCGCTACTACAGAGGACGCTGCCAGACA 665
Db 301 GTGGTGTGCCCCCACCACCCCTTCTCGATCGCTACTACAGAGGACGCTGCCAGACA 360
Qy 666 TACCGGGACCTCCGGAAGGGTGTATGTATGTCCTTACACCCAGGCGAAGTGGGAAGGGAG 725
Db 361 TACCGGGACCTCCGGAAGGGTGTATGTATGTCCTTACACCCAGGCGAAGTGGGAAGGGAG 420
Qy 726 CTGGCAGCGACCTGTGAAGCATCCCCATGGCCCAACGCTCACTGTGGTGGCCACAT 785
Db 421 CTGGCAGCGACCTGTGAAGCATCCCCATGGCCCAACGCTCACTGTGGTGGCCACAT 480
Qy 786 GCTGCCATCACTGAATCAGACAAGTCTTTCATCAACGCTCCAACTGGGAAGGCAATCCG 845
Db 481 GCTGCCATCACTGAATCAGACAAGTCTTTCATCAACGCTCCAACTGGGAAGGCAATCCG 540
Qy 846 GGGTGGCCCTATGCTGAGATGGCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 905
Db 541 GGGTGGCCCTATGCTGAGATGGCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 600
Qy 906 CTGGTAAAGCAGACCCACGCTCCCAACCTCTTCTCCCTGCGAGCTTTGGTGGCTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACGCTCCCAACCTCTTCTCCCTGCGAGCTTTGGTGGCTGGCTTC 660
Qy 966 CCCCTCAACGACTCTGAAGTGTGGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCCTCAACGACTCTGAAGTGTGGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 720
Qy 1026 GACCACTCGCTACACAGGCACTCTGTGTATACACCCATCCCGGGGAGTGGTATAT 1085
Db 721 GACCACTCGCTACACAGGCACTCTGTGTATACACCCATCCCGGGGAGTGGTATAT 780
Qy 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGATCTGAATGGACTGCAAGGAG 1145
Db 781 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGATCTGAATGGACTGCAAGGAG 840
Qy 1146 TACAATATGACAAGAGCATTTGTGGAGTGGACACCAACCTTCGTTTGGCCCAAGAAA 1205
Db 841 TACAATATGACAAGAGCATTTGTGGAGTGGACACCAACCTTCGTTTGGCCCAAGAAA 900
Qy 1206 GTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCCTCTCCACGGAGAAGTTCCTGTAT 1265
Db 901 GTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCCTCTCCACGGAGAAGTTCCTGTAT 960
Qy 1266 GTTCTTGGCTAGGAGCAGCTGGTGTGCTGGAAGCAGGACCCACCTTCGTTGGAACTT 1325
Db 961 GTTCTTGGCTAGGAGCAGCTGGTGTGCTGGAAGCAGGACCCACCTTCGTTGGAACTT 1020
Qy 1326 TTCCAGTCACTCACTTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCAATCAC 1385
Db 1021 TTCCAGTCACTCACTTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCAATCAC 1080
Qy 1386 ATCTTCCGACCAATACCTCGGCCAGTGGAGATGTGGCCACGCTCCCAAGACGACTGT 1445
Db 1081 ATCTTCCGACCAATACCTCGGCCAGTGGAGATGTGGCCACGCTCCCAAGACGACTGT 1140
Qy 1446 TACAAGTTGGCATCTCAGATCATCCACGGGCACTCTTATGGAGCTGTATCATGGAG 1505
Db 1141 TACAAGTTGGCATCTCAGATCATCCACGGGCACTGTATGGAGCTGTATCATGGAG 1200
Qy 1506 GGCTTCTACGTTGTCTTTCATCGGGCCGAAAGCAATTTGCTGTACGCGCTTCC 1565
Db 1201 GGCTTCTACGTTGTCTTTCATCGGGCCGAAAGCAATTTGCTGTACGCGCTTCC 1260
Qy 1566 CATGTGCAGGATGAGTTGAGGAGCGGAGCGGTGGAGGCCCTTTGTCACTTGGACATG 1625
Db 1261 CATGTGCAGGATGAGTTGAGGAGCGGAGCGGTGGAGGCCCTTTGTCACTTGGACATG 1320
Qy 1626 GAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCTAT 1685
Db 1321 GAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCTAT 1380

Qy 1686 GTCATGGCTCCCATCTCGGCCCTCTTTCATGCTGGCACCTCTGCCTCATGCTGTGCTCACTGG 1745
Db 1381 GTCATGGCTCCCATCTCGGCCCTCTTTCATGCTGGCACCTCTGCCTCATGCTGTGCTCACTGG 1440
Qy 1746 CGCTGGCTCCGCTGGCTGGCCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGGCTCCGCTGGCTGGCCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTG 1500
Qy 1806 AAGTGAAGGAGGCCATGGGCGAGAGATAGAGATTTCCCTGGACACACCTCCCTGGCTTCA 1865
Db 1501 AAGTGAAGGAGGCCATGGGCGAGAGATAGAGATTTCCCTGGACACACCTCCCTGGCTTCA 1560
Qy 1866 CTTTGGTCAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGACCTTCAGGACCTCC 1925
Db 1561 CTTTGGTCAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGACCTTCAGGACCTCC 1620
Qy 1926 CCACCCACCAATGGCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGGTGGGTTCAG 1985
Db 1621 CCACCCACCAATGGCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGGTGGGTTCAG 1680
Qy 1986 GGAAGTGTACCTGTAGGAAACAGAAAGAGAAAGAGCACTCTGCTGGCGGGAATACT 2045
Db 1681 GGAAGTGTACCTGTAGGAAACAGAAAGAGAAAGAGCACTCTGCTGGCGGGAATACT 1740
Qy 2046 CTTGGTCACTCAAAATTTAAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 2105
Db 1741 CTTGGTCACTCAAAATTTAAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 1800
Qy 2106 TTGTCACCACTTCCTTTAAATTTCTCAACCCAAAGTATTTCTTTCTTTAGTTTTCAGAA 2165
Db 1801 TTGTCACCACTTCCTTTAAATTTCTCAACCCAAAGTATTTCTTTCTTTAGTTTTCAGAA 1860
Qy 2166 GTACTGGCATCACACGAGCTTACCTTGGCGTGTCTCCCTGTGGTACCCCTGGCAGAGAG 2225
Db 1861 GTACTGGCATCACACGAGCTTACCTTGGCGTGTCTCCCTGTGGTACCCCTGGCAGAGAG 1920
Qy 2226 AGACCAAGCTTGTTCCTCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGTCTATT 2285
Db 1921 AGACCAAGCTTGTTCCTCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGTCTATT 1980
Qy 2286 TGCTTTAGAGACGAGGACTGTATAACCAAGCCTAACATTTGGTGAAGATTGCCTCTTGA 2345
Db 1981 TGCTTTAGAGACGAGGACTGTATAACCAAGCCTAACATTTGGTGAAGATTGCCTCTTGA 2040
Qy 2346 ATT 2348
Db 2041 ATT 2043

RESULT 2

US-09-548-367D-3
: Sequence 3, Application US/09548367D
: Patent No. 6440698
: GENERAL INFORMATION:
: APPLICANT: GURNEY ET AL.
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: 29915/6280H
: CURRENT APPLICATION NUMBER: US/09/548,367D
: CURRENT FILING DATE: 2000-04-12
: PRIOR APPLICATION NUMBER: US 60/155,493
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: US 09/404,133
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: PCT/US99/20881
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: US 60/101,594
: PRIOR FILING DATE: 1998-09-24
: NUMBER OF SEQ ID NOS: 73
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 2070
: TYPE: DNA

RESULT 3

US-09-009-191-1

; Sequence 1, Application US/09009191

; Patent No. 6319689

; GENERAL INFORMATION:

; APPLICANT: POWELL, DAVID

; APPLICANT: CHAPMAN, CONRAD

; APPLICANT: MURPHY, KAY

; APPLICANT: SMITH, TRUDI

; TITLE OF INVENTION: ASP2

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/009,191

; FILING DATE: 20-JAN-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: UK 9701684.4

; FILING DATE: 28-JAN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F.

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-70368

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2541 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-09-009-191-1

Query Match 86.2%; Score 2024; DB 4; Length 2541;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2038; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 306 ATGGCCCAAGCCCTGCCCTGGCTCTGCTGTGATGGCGGGGAGTGCTGCTGCCCCAC 365

Db 1 ATGGCCCAAGCCCTGCCCTGGCTCTGCTGTGATGGCGGGGAGTGCTGCTGCCCCAC 60

QY 366 GGCACCCAGCAGGCGATCGGCTGCCCTCCGCGAGCGCGCTGGGGGGCGCCCTGGGG 425

Db 61 GGCACCCAGCAGGCGATCGGCTGCCCTCCGCGAGCGCGCTGGGGGGCGCCCTGGGG 120

QY 426 CTGGCGCTGCCCGGAGACCCAGCAGCGCGCGAGGCGCGCGCGCGGAGGCGAGCTTT 485

Db 121 CTGGCGCTGCCCGGAGACCCAGCAGCGCGCGAGGCGCGCGCGCGGAGGCGAGCTTT 180

QY 486 GTGAGATGTGGGACAACCTGAGGGCAAGTGGGGCAGGCTACTAGTGGAGATGACC 545

Db 181 GTGAGATGTGGGACAACCTGAGGGCAAGTGGGGCAGGCTACTAGTGGAGATGACC 240

QY 546 GTGGCAGCCCCCGCAGACGCTCAACATCTGTGGATACAGCAGCAGTAACCTTGCA 605

Db 241 GTGGCAGCCCCCGCAGACGCTCAACATCTGTGGATACAGCAGCAGTAACCTTGCA 300

QY 606 GTGGTGCTGCCCCCACCACCCCTTCTGCTGCTACTACAGAGCAGCTGTCCAGCACA 665
Db 301 GTGGTGCTGCCCCCACCACCCCTTCTGCTGCTACTACAGAGCAGCTGTCCAGCACA 360
QY 666 TACCGGGACCTCCGGGAAGGGTGTATGTGCCCTTACACCCAGGCGAAGTGGGAAGGGAG 725
Db 361 TACCGGGACCTCCGGGAAGGGTGTATGTAGCCCTTACACCCAGGCGAAGTGGGAAGGGAG 420
QY 726 CTGGGACCGGAGCTGTAAGCATCCCATGGCCCAACGTCACCTGTGGTGGTCCACATTT 785
Db 421 CTGGGACCGGAGCTGTAAGCATCCCATGGCCCAACGTCACCTGTGGTGGTCCACATTT 480
QY 786 GCTGCCATCAGTGAATCAGACAAAGTTCTTTCATCAACGGCTCCAACTGGGAAGGCACTCTG 845
Db 481 GCTGCCATCAGTGAATCAGACAAAGTTCTTTCATCAACGGCTCCAACTGGGAAGGCACTCTG 540
QY 846 GGGTGGGCTATGCTGAGATTGCCAGGCTTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 905
Db 541 GGGTGGGCTATGCTGAGATTGCCAGGCTTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 600
QY 906 CTGGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTGCGAGCTTTGTGGTCTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTGCGAGCTTTGTGGTCTGGCTTC 660
QY 966 CCCCTCAACCACTCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCCTCAACCACTCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACTCGCTGTACACAGCAGTCTCTGTGTATACCCATCCCGGGGAGTGGTATTTAT 1085
Db 721 GACCACTCGCTGTACACAGCAGTCTCTGTGTATACCCATCCCGGGGAGTGGTATTTAT 780
QY 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGATCTGAAATGACATGCAAGGAG 1145
Db 781 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGATCTGAAATGACATGCAAGGAG 840
QY 1146 TACAACTATGACAAGAGCATTTGTGACAGTGGCAGCACCAACCTTTCGTTGGCCCAAGAA 1205
Db 841 TACAACTATGACAAGAGCATTTGTGACAGTGGCAGCACCAACCTTTCGTTGGCCCAAGAA 900
QY 1206 GTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCTTCTCCAGGAGAGTTCCCTGAT 1265
Db 901 GTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCTTCTCCAGGAGAGTTCCCTGAT 960
QY 1266 GGTTCCTGGCTAGGAGCAGCTGCTGTGCTGGAAGCAGCAGCACCCCTTTGGAACTT 1325
Db 961 GGTTCCTGGCTAGGAGCAGCTGCTGTGCTGGAAGCAGCAGCACCCCTTTGGAACTT 1020
QY 1326 TTCCAGTCACTCACTCTACCTTAATGGGTGAGGTACCAACAGTCTCTCCGCACTACC 1385
Db 1021 TTCCAGTCACTCACTCTACCTTAATGGGTGAGGTACCAACAGTCTCTCCGCACTACC 1080
QY 1386 ATCTTCCGACAGCAATACCTGCGGCGCAGTGGAAAGATGTGGGCACTCCCAAGCAGCTGT 1445
Db 1081 ATCTTCCGACAGCAATACCTGCGGCGCAGTGGAAAGATGTGGGCACTCCCAAGCAGCTGT 1140
QY 1446 TACAAGTTTGGCATCTCAGTCACTCACCGGCACTGTTATGGAGCTGTTTATCATCGAG 1505
Db 1141 TACAAGTTTGGCATCTCAGTCACTCACCGGCACTGTTATGGAGCTGTTTATCATCGAG 1200
QY 1506 GGTTCCTACGTTCTCTTTGATCGGGCCGAAAGCAATTTGGCTTTGCTCAGCGCTTCG 1565
Db 1201 GGTTCCTACGTTCTCTTTGATCGGGCCGAAAGCAATTTGGCTTTGCTCAGCGCTTCG 1260
QY 1566 CATGTGACGAGTGTGATTCAGGAGCGGAGCGTGAAGGCCCTTTTGTACCTTTGGACATG 1625
Db 1261 CATGTGACGAGTGTGATTCAGGAGCGGAGCGTGAAGGCCCTTTTGTACCTTTGGACATG 1320
QY 1626 GAAGACTGTGGCTTACAACATTTCCACACAGATGAGTCAAGCTCAACCTCATGACCATAGCTTAT 1685
Db 1321 GAAGACTGTGGCTTACAACATTTCCACACAGATGAGTCAAGCTCAACCTCATGACCATAGCTTAT 1380
QY 1686 GTCATGGCTGCCATCTGCGCCCTCTTTATGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG 1745

Db 567 -----||||| 585
QY 966 CCCCTCAACAGTCTGAAGTCTGGCCTCTGTCGGAGGAGCATGATCATTTGGAGGTATC 1025
Db 586 CCCCTCAACAGTCTGAAGTCTGGCCTCTGTCGGAGGAGCATGATCATTTGGAGGTATC 645
QY 1026 GACCACTCGCTGTACACAGCAGTCTCTGGTATACACCCATCCCGCGGGAGTGGTATTAT 1085
Db 646 GACCACTCGCTGTACACAGCAGTCTCTGGTATACACCCATCCCGCGGGAGTGGTATTAT 705
QY 1086 GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGATCTGAATAATGGACTGCAAGGAG 1145
Db 706 GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGATCTGAATAATGGACTGCAAGGAG 765
QY 1146 TACAACATATGACAAGAGCATTGTGGACAGTGGCACACCAACCTTCCTTTGCCCAAGAAA 1205
Db 766 TACAACATATGACAAGAGCATTGTGGACAGTGGCACACCAACCTTCCTTTGCCCAAGAAA 825
QY 1206 GTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCCTCCTCCAGGAGAAAGTTCCCTGAT 1265
Db 826 GTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCCTCCTCCAGGAGAAAGTTCCCTGAT 885
QY 1266 GGTTCCTGGCTAGGAGAGCAGTGGTGTGGCAAGCAGCACCCCTTGGAAACATT 1325
Db 886 GGTTCCTGGCTAGGAGAGCAGTGGTGTGGCAAGCAGCACCCCTTGGAAACATT 945
QY 1326 TTCCCACTCATCTCACTCTACCTAATGGTTCAGTTTACCAACCACTCTTCCGCATCACC 1385
Db 946 TTCCCACTCATCTCACTCTACCTAATGGTTCAGTTTACCAACCACTCTTCCGCATCACC 1005
QY 1386 ATCTTTCCGAGCAATACCTCGCGCCAGTGGAGATGTGGCCAGTCCCAAGACGACTGT 1445
Db 1006 ATCTTTCCGAGCAATACCTCGCGCCAGTGGAGATGTGGCCAGTCCCAAGACGACTGT 1065
QY 1446 TACAAGTTTGGCATCTCACAGTATCATCCAGCGGCACTGTTATGGAGGTGTTATCATGGAG 1505
Db 1066 TACAAGTTTGGCATCTCACAGTATCATCCAGCGGCACTGTTATGGAGGTGTTATCATGGAG 1125
QY 1506 GGCTTCTACGTTGTCTTGTATCGGGCCGGAAGACGAATGGCTTGTCTCAGCGCTTGC 1565
Db 1126 GGCTTCTACGTTGTCTTGTATCGGGCCGGAAGACGAATGGCTTGTCTCAGCGCTTGC 1185
QY 1566 CATGTGCAGCATGATTTCCAGGACGCGAGCGTGGAGGCCCTTTTGTACCTTGGACATG 1625
Db 1186 CATGTGCAGCATGATTTCCAGGACGCGAGCGTGGAGGCCCTTTTGTACCTTGGACATG 1245
QY 1626 GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
Db 1246 GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1305
QY 1686 GTCATGGCTGCCATCTCGGCCCTCTTCACTGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG 1745
Db 1306 GTCATGGCTGCCATCTCGGCCCTCTTCACTGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG 1365
QY 1746 CGCTGCCCTCCGCTCGCTCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTCGTG 1805
Db 1366 CGCTGCCCTCCGCTCGCTCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTCGTG 1425
QY 1806 AAGTGAGAGGCCCATGGGAGAGATAGAGATTTCCCTTGGACACACACCTCCCGTGGTTCA 1865
Db 1426 AAGTGAGAGGCCCATGGGAGAGATAGAGATTTCCCTTGGACACACACCTCCCGTGGTTCA 1485
QY 1866 CTTTGGTTCACAAGTGAAGACACAGATGGCACCTGTGGCCAGACACCTCAGGACCCCTCC 1925
Db 1486 CTTTGGTTCACAAGTGAAGACACAGATGGCACCTGTGGCCAGACACCTCAGGACCCCTCC 1545
QY 1926 CCACCCACCAATGCCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGTGGGTTCAG 1985
Db 1546 CCACCCACCAATGCCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGTGGGTTCAG 1605
QY 1986 GGACTGTACCTGTAGGAACAGAAAGAGAGAAAGACACTCTGCTGGCGGGAATACT 2045
|||||

Db 1606 GGACTGTACCTGTAGGAACAGAAAGAGAAAGACACTCTGCTGGCGGAATACT 1665
QY 2046 CTTGGTCCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 2105
Db 1666 CTTGGTCCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 1725
QY 2106 TTGTCCACCATTCTCTTAAATTTCTCCAAACCCAAAGTATTTCTTTCTTTAGTTTCAGAA 2165
Db 1726 TTGTCCACCATTCTCTTAAATTTCTCCAAACCCAAAGTATTTCTTTCTTTAGTTTCAGAA 1785
QY 2166 GTACTGCGATCACACGAGGTTACCTTTGGCGTGTCTCCCTGCTGCTACCTGGCAGAGAG 2225
Db 1786 GTACTGCGATCACACGAGGTTACCTTTGGCGTGTCTCCCTGCTGCTACCTGGCAGAGAG 1845
QY 2226 AGACCAAGCTTGTTCCTCTGCTGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTATT 2285
Db 1846 AGACCAAGCTTGTTCCTCTGCTGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTATT 1905
QY 2286 TGCTTTAGAGACAGGGACTGTATAAACAAGCCTTAACATTTGGTGCAAAAGATTTGCCTCTTGA 2345
Db 1906 TGCTTTAGAGACAGGGACTGTATAAACAAGCCTTAACATTTGGTGCAAAAGATTTGCCTCTTGA 1965
QY 2346 A 2346
Db 1966 A 1966

RESULT 6
US-09-009-191-3
; Sequence 3, Application US/09009191
; Patent No. 6319689
; GENERAL INFORMATION:
; APPLICANT: POWELL, DAVID
; APPLICANT: CHAPMAN, CONRAD
; APPLICANT: MURPHY, KAY
; APPLICANT: SMITH, TRUDI
; TITLE OF INVENTION: ASP2
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,191
; FILING DATE: 20-JAN-1998
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: UK 9701684.4
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70368
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-009-191-3

Query Match	78.2%;	Score 1835.4;	DB 4;	Length 2370;		
Best Local Similarity	99.1%;	Prod. No. 0;				
Matches 1856;	Conservative	0;	Mismatches	16; Indels	1; Gaps	1;
QY	477	GGCAGCTTTGTCGAGATGTTGGACGACCACTGAGGGCAAGTCGGGCGAGGCTACTACGCTG	536			
Db	1	GGCAGCTTTGTGGAGATGGTGGACAACTGAGGGCAAGTCGGGCGAGGCTACTACGCTG	60			
QY	537	GAGATGACCGTGGGGAGCCCCCGCAGACGCTCAACATCCTTGGTGGATACAGGCAGCAGT	596			
Db	61	GAGATGACCGTGGGAGCCCCCGCAGACGCTCAACATCCTTGGTGGATACAGGCAGCAGT	120			
QY	597	AACTTTGCAGTGGGTGCTGCCGCCACCCCTTCTCGATCGCTACTACAGAGGCGAGCTG	656			
Db	121	AACTTTGCAGTGGGTGCTGCCGCCACCCCTTCTCGATCGCTACTACAGAGGCGAGCTG	180			
QY	657	TCCAGCACATACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGG	716			
Db	181	TCCAGCACATACCGGGACCTCCGGAAGGGTGTGTATGAGCCCTACACCCAGGGCAAGTGG	240			
QY	717	GAAGGGAGCTGGGACACCGACTGGTTAAGCATCCCCCATGGCCCCAACGTCACCTGTGCGGT	776			
Db	241	GAAGGGAGCTGGGACACCGACTGGTTAAGCATCCCCCATGGCCCCAACGTCACCTGTGCGGT	300			
QY	777	GCCAACTTGTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAA	836			
Db	301	GCCAACTTGTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAA	360			
QY	837	GGCATCTCGGGCTGGCCCTATGTGAGATTGCCAGGCTCGAGACTCCCTGGAGCCCTTC	896			
Db	361	GGCATCTCGGGCTGGCCCTATGTGAGATTGCCAGGCTCGAGACTCCCTGGAGCCCTTC	420			
QY	897	TTTGACTCTCTGTTAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGACGCTTTGTGGT	956			
Db	421	TTTGACTCTCTGTTAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGACGCTTTGTGGT	480			
QY	957	GCTGSCCTTCCCCCTCAACAGCTCTGAAGTGTGTGCCCTCTGTGCGAGGGAGCATGATCATTT	1016			
Db	481	GCTGSCCTTCCCCCTCAACAGCTCTGAAGTGTGTGCCCTCTGTGCGAGGGAGCATGATCATTT	540			
QY	1017	GGAGGTATCGACCACTCGCTGTACACAGGAGCTCTCTGGTATACACCCATCGGCGGGAG	1076			
Db	541	GGAGGTATCGACCACTCGCTGTACACAGGAGCTCTCTGGTATACACCCATCGGCGGGAG	600			
QY	1077	TGGTATTATGAGGTGATCATTTGTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGAC	1136			
Db	601	TGGTATTATGAGGTGATCATTTGTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGAC	660			
QY	1137	TGCAAGGAGTACAATATGACAAGAGCATTTGTGGACAGTGGCACCAACCTTCGTTTG	1196			
Db	661	TGCAAGGAGTACAATATGACAAGAGCATTTGTGGACAGTGGCACCAACCTTCGTTTG	720			
QY	1197	CCCAAGAAGTCTTTGAAGCTGCAGTCAAAATCCATCAGGAGCGCTCTCCACGGAGCAAG	1256			
Db	721	CCCAAGAAGTCTTTGAAGCTGCAGTCAAAATCCATCAGGAGCGCTCTCCACGGAGCAAG	780			
QY	1257	TTCCCTGATGTTTCTCGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACACCCCT	1316			
Db	781	TTCCCTGATGTTTCTCGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACACCCCT	840			
QY	1317	TGGAACATTTTCCCACTCATCTCACTACCTAATGGGTGAGGTTACCAACAGTCCCTTC	1376			
Db	841	TGGAACATTTTCCCACTCATCTCACTACCTAATGGGTGAGGTTACCAACAGTCCCTTC	900			
QY	1377	CGCATCACCATCTCTCCGAGCAATACCTGCGGGCAGTGGGAAGATGTGCCACAGTCCCAA	1436			
Db	901	CGCATCACCATCTCTCCGAGCAATACCTGCGGGCAGTGGGAAGATGTGCCACAGTCCCAA	960			
QY	1437	GACGACTGTTACAAGTTTGCCATCTCACAGTCAATCCACGGGCACTGTTATGAGGAGCTGTT	1496			
Db	961	GACGACTGTTACAAGTTTGCCATCTCACAGTCAATCCACGGGCACTGTTATGAGGAGCTGTT	1020			

RESULT 7

RESULT /
IIS-09-548-372D-7

US-09-548-372D-7
: Sequence 7 Application IIS/09548372D

; Sequence 7, Application No. 6420524

; Patent No. 6420534

; GENERAL INFORMATION:

APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE: A REVIEW OF THE LITERATURE.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SE

; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 29915/62801

; FILE REFERENCE: 29915/6280I
 ; CURRENT APPLICATION NUMBER: US/09/548 372D

; CURRENT APPLICATION NUMBER: US/09/548,313

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

[illegible]

Db 1501 AAGTAAGGAGGCTCGTGGGAGAGATGAGAGACGCCCTGGACACATCTGGGTGGTTC 1560
Qy 1866 CTTTGGTGCACA--AGTAGGAGACACAGATGGACCTGTGGCCAGAGCACCTCAGGACCT 1923
Db 1561 CTTTGGTGCATAGTGTGGAGCTATGATGGTACCTGTGGCCAGAGCACCTCAGGACCT 1620
Qy 1924 CCCACCCACCAATAGCTCTGCTTGTGAGAGAAAGGCTGGCAAGTGGGTTC 1983
Db 1621 CACCAACTGCCAATGCTTCTGGGTGACAGA-ACAGAGAAATCAGGCAAGCTGATTAC 1679
Qy 1984 AGGACTGTACCTGTAGGAACAGAAAGAGAGAAAGACACTGTGCTGGGGGGAATA 2043
Db 1680 AGGCTTTGCACCTGTAGGACACAGGAGGAGGAGGAGGAGCGCTTCTGGTGGCAGGAATA 1739
Qy 2044 CTCTGTGTACCTTCAATTTAAAGTCGGGAATTTCTGCTGTGAACTTTCAGCCCTGAAC 2103
Db 1740 TCCTTAGGCAACCAAACTTGAGT-TGGAATTTTGTCTTGAAGCTTCAGCCCTGACC 1798
Qy 2104 CTTTGTCCAGCATTCCTTTAAATTTCCAAACCAAGTATTTCTTTTCTTAGTTTCAG 2163
Db 1799 CTCTGCCAGCA-TCCTTTAGAGTCTCCAACCTAAAGTATTTATATGTC---CTTCCAG 1854
Qy 2164 AAGTACTGGCATCACAGCGAGTTACCTTGGCGTGTCTCCCTGTGTGACCTGGCAGAGA 2223
Db 1855 AAGTACTGGCGTCTACTCAGGCTACC-CGGCATGTCTCCCTGTGTGTACCTGGCAGAGA 1913
Qy 2224 AGAGACAAAGCTGTTTCCCTGTGGCCAAAGTCAGTAGGAGAGATGACAGATTTCCTA 2283
Db 1914 AAGGCCAAATC-TCATTTCCCTGTGGCCAAAGTCAGCAGAGAAGGTG--AAGTTTCCCA 1970
Qy 2284 TTTGCTTTAGACAGAGGAGCTGTATAACAAAGCCTAACATTTGGTGCAAGATTGCCCTTT 2343
Db 1971 GTTGTCTTAGTAGGAGCTGCAGACTCAAGCT-ACACTGGTACAAAGACTGCGTCTT 2029
Qy 2344 GA 2345
Db 2030 GA 2031

RESULT 9

US-09-548-372D-29
; Sequence 29, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; PRIOR FILING DATE: 2000-04-12
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-372D-29

Query Match 57.7%; Score 1355.8; DB 4; Length 1362;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1357; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 306 ATGCCCAAGCCCTGCCTGCTGCTGTGGATGGCGGGGAGTGTGCGCTGCCAC 365
|||||

Db 1 ATGCCCAAGCCCTGCCTGGCTCTCTGCTGTGGATGGCGGGGAGTGTCTGCCTGCCAC 60
Qy 366 GGCACCCAGACGCGATCCGGCTCGCCCTCGGCAGCGCCCTGGGGGGCGCCCTGGGG 425
Db 61 GGCACCCAGACGCGATCCGGCTCGCCCTCGGCAGCGCCCTGGGGGGCGCCCTGGGG 120
Qy 426 CTGGGCTGCCCCGGGAGACCGAGAGAGCCGAGGAGCCCGGCGGAGGCGGAGCTTT 485
Db 121 CTGGGCTGCCCCGGGAGACCGAGAGAGCCGAGGAGCCCGGCGGAGGCGGAGCTTT 180
Qy 486 GTGAGATGTGGACAACCTGTAGGGGCAAGTGTGGGGAGGCTACTAGTGGAGATGACC 545
Db 181 GTGAGATGTGGACAACCTGTAGGGGCAAGTGTGGGGAGGCTACTAGTGGAGATGACC 240
Qy 546 GTGGGAGCCCCCGCAGACGCTCAACATCTGTGTGATACAGCAGCAGTAACCTTGCA 605
Db 241 GTGGGAGCCCCCGCAGACGCTCAACATCTGTGTGATACAGCAGCAGTAACCTTGCA 300
Qy 606 GTGGTGTCTGCCCGCCACCCCTTCTCTGCATCGCTACTACCAGAGGACGCTTCCAGCACA 665
Db 301 GTGGTGTCTGCCCGCCACCCCTTCTCTGCATCGCTACTACCAGAGGACGCTTCCAGCACA 360
Qy 666 TACCGGACCTCCGGAAGGCTGTATGTGCCCTTACACCCAGGCAAGTGGGAAGGGAG 725
Db 361 TACCGGACCTCCGGAAGGCTGTATGTGCCCTTACACCCAGGCAAGTGGGAAGGGAG 420
Qy 726 CTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCCAACGCTACTGCTGCCAACAATT 785
Db 421 CTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCCAACGCTACTGCTGCCAACAATT 480
Qy 786 GTCGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCTCTG 845
Db 481 GTCGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCTCTG 540
Qy 846 GGGCTGCCCTATGCTGAGATTTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 905
Db 541 GGGCTGCCCTATGCTGAGATTTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 600
Qy 906 CTGTAAGCAGACCCACGTTTCCCAACCTCTCTCCCTGCGAGCTTTGTTGGTGGCTTC 965
Db 601 CTGTAAGCAGACCCACGTTTCCCAACCTCTCTCCCTGCGAGCTTTGTTGGTGGCTTC 660
Qy 966 CCCCTCAACGCTCTGAAGTCTGGCCTCTGTGCGAGGAGCAGTATCATTTGGAGGTATC 1025
Db 661 CCCCTCAACGCTCTGAAGTCTGGCCTCTGTGCGAGGAGCAGTATCATTTGGAGGTATC 720
Qy 1026 GACCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGGTATTAT 1085
Db 721 GACCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGGTATTAT 780
Qy 1086 GAGTGTATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 1145
Db 781 GAGTGTATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 840
Qy 1146 TACAATATGACAAGACGATTTGGACAGTGGCAGCACCACCAACCTTCGTTCCCAAGAAA 1205
Db 841 TACAATATGACAAGACGATTTGGACAGTGGCAGCACCACCAACCTTCGTTCCCAAGAAA 900
Qy 1206 GTGTTTCAAGCTGCAGTCAAAATCCATCAAGCAGCCTCTCCAGGAGAAAGTTCCTGAT 1265
Db 901 GTGTTTCAAGCTGCAGTCAAAATCCATCAAGCAGCCTCTCCAGGAGAAAGTTCCTGAT 960
Qy 1266 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGCGCAAGCAGCAGCACCACCTTTGGAACATT 1325
Db 961 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGCGCAAGCAGCAGCACCACCTTTGGAACATT 1020
Qy 1326 TTCCAGTCACTCACTCTACCTAATGGTGAGTTTACCAACCAAGTCTCTCCGATCACC 1385
Db 1021 TTCCAGTCACTCACTCTACCTAATGGTGAGTTTACCAACCAAGTCTCTCCGATCACC 1080
Qy 1386 ATCCTTCCGAGCAATACCTGCGGCCAGTGGAAAGATGTGGCCACGTTCCCAAGAGACTGT 1445
Db 1081 ATCCTTCCGAGCAATACCTGCGGCCAGTGGAAAGATGTGGCCAGTGTCCCAAGAGACTGT 1140

Qy	1446	TACAAAGTTTGCCATCTCAGATCCACGGGCACCTGTTATGGGAGCTGTTATCATGGAG	1505
Db	1441	TACAAGTTTGCCATCTCAGATCCACGGGCACCTGTTATGGGAGCTGTTATCATGGAG	1200
Qy	1506	GGCTTCTACGTTGCTTTTGATCGGGCCGCAAAACGAATTTGGCTGCTGTCACGCGCTTGC	1565
Db	1201	GGCTTCTACGTTGCTTTTGATCGGGCCGCAAAACGAATTTGGCTGCTGTCACGCGCTTGC	1260
Qy	1566	CATGTGCACGATGAGTTCAGGACGCGGTTGGAAGGCCCTTTTGTCACTTTGGACATG	1625
Db	1281	CATGTGCACGATGAGTTCAGGACGCGGTTGGAAGGCCCTTTTGTCACTTTGGACATG	1320
Qy	1626	GAAGACTCTGCTACAAACATTCACAGACAGATGAGTCA	1664
Db	1321	GAAGACTCTGCTACAAACATTCACAGACAGATGAGTCA	1359
RESULT 10			
US-09-548-367D-29			
: Sequence 29, Application US/09548367D			
: Patent No. 6440698			
: GENERAL INFORMATION:			
: APPLICANT: GURNEY ET AL.			
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR			
: FILE OF INVENTION: THEREOF			
: FILE REFERENCE: 29915/6280H			
: CURRENT APPLICATION NUMBER: US/09/548,367D			
: CURRENT FILING DATE: 2000-04-12			
: PRIOR APPLICATION NUMBER: US 60/155,493			
: PRIOR FILING DATE: 1999-09-23			
: PRIOR APPLICATION NUMBER: US 09/404,133			
: PRIOR FILING DATE: 1999-09-23			
: PRIOR APPLICATION NUMBER: PCT/US99/20881			
: PRIOR FILING DATE: 1999-09-23			
: PRIOR APPLICATION NUMBER: US 60/101,594			
: PRIOR FILING DATE: 1998-09-24			
: NUMBER OF SEQ ID NOS: 73			
: SOFTWARE: PatentIn version 3.1			
: SEQ ID NO 29			
: LENGTH: 1362			
: TYPE: DNA			
: ORGANISM: Homo sapiens			
US-09-548-367D-29			

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT II
US-09-548-372D-31

03 09 348 372D 31
: Sequence 31, Application US/09548372D

Patent No. 6420534

Qy	1326	TTCCAGTCATCTCACTCTACCTAATGGGTGAGGTTTACCAACGACGTCCTTTCCGCATCACC	1386
Db	1021	TTCCAGTCATCTCACTCTACCTCTACCTAATGGGTGAGGTTTACCAACGACGTCCTTTCCGCATCACC	1080
Qy	1386	ATCCTTCCGCAGCAATACCTCGGGCCAGTGGGAAGATGTGGCCACACGTCCTCCCAAGACGACTGT	1445
Db	1081	ATCCTTCCGCAGCAATACCTCGGGCCAGTGGGAAGATGTGGCCACGTCCTCCCAAGACGACTGT	1140
Qy	1446	TACAAGTTTGGCATCTCACAGTCATCCACGGGCACCTGTTTATGGGAGCTGTTCATCATGGAG	1505
Db	1141	TACAAGTTTGGCATCTCACAGTCATCCACGGGCACCTGTTTATGGGAGCTGTTCATCATGGAG	1200
Qy	1506	GGCTTCTACGTTGCTCTTGTATCGGGCCCGCAAAACGAATTTGGCTTCTGTCTCAGCGCTTGC	1565
Db	1201	GGCTTCTACGTTGCTCTTGTATCGGGCCCGCAAAACGAATTTGGCTTCTGTCTCAGCGCTTGC	1260
Qy	1566	CATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAAGGCCCTTTTGTCACTTTGGACATG	1625
Db	1261	CATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAAGGCCCTTTTGTCACTTTGGACATG	1320
Qy	1626	GAAGACTGTGCTTACAACATTTCCACAGACAGATGAGTCA	1664
Db	1321	GAAGACTGTGCTTACAACATTTCCACAGACAGATGAGTCA	1359

RESULT 13

US-09-713-158-1

; Sequence 1, Application US/09713158
; Patent No. 6361975
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: LI, XIAOTONG
; APPLICANT: POWELL, DAVID J.
; APPLICANT: CHRISTIE, GARY
; TITLE OF INVENTION: MOUSE ASPARTIC SECRETASE-2 (NASP-2)
; FILE REFERENCE: GP-70660
; CURRENT APPLICATION NUMBER: US/09/713,158
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 60/165,800
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: MUS MUSCULUS
US-09-713-158-1

Db 301 GTGGGGCTGCCACACACCCCTTCTGCTGCTACTTACCAGAGGAGCTGTCCAGCACA 360
Qy TACCGGACCTCGGAGAGGTGTATGTGCTTACACCCAGGACAGTGGGAGGGAG 725
Db 361 TATCGAGACCTCCGAAGGGGTGTATGTGCTTACACCCAGGACAGTGGGAGGGAA 420
Qy 726 CTGGGACCGACCTGGTGAAGCATCCCCATGGCCCCAACGTCATGTGCGTGGCAACATT 785
Db 421 CTGGGACCGACCTGGTGAAGCATCCCTCATGCGCCCCAACGTCATGTGCGTGGCAACATT 480
Qy 786 GTGGCANTACTGAATCAGACAAAGTTCTTCATCAAGGCTTCCAACTGGGAGGATCCTG 845
Db 481 GCTGCCATCACTGAATCGGACAAAGTTCTTCATCAATGGTTCCAACTGGGAGGCACTCCTA 540
Qy 846 GGCTGGCTATCTGATGATTCAGGCTTCAGGCTTCAGGACTCCCTGGAGCTTTCTTGACTCT 905
Db 541 GGCTGGCTATCTGATGATTCAGGCTTCAGGCTTCAGGACTTCCTTGAGGCTTTCTTGACTCC 600
Qy 906 CTGGTAAAGCAGACCCACAGCTTCCCAACCTCTTCTCCTGCGAGCTTTGTGCTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACAGCTTCCCAACCTCTTCTCCTGCGAGCTTCCTGGCTGGCTTC 660
Qy 966 CCCCTCAACAGTCTGAAGTGTGGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCCTCAACAGCAGCCAGGCACTGGGCTCGGTGGGAGGAGCATGATCATTTGGTGTATC 720
Qy 1026 GACCACTCGTGTACAGAGCAGTCTGCTGTATACACCCATCGCGGGAGTGGTATTAT 1085
Db 721 GACCACTCGTGTATACAGGCGCATCTCTGCTGATACACCCATCGCGGGAGTGGTATTAT 780
Qy 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGATCTGAAATGGACTGCAAGGAG 1145
Db 781 GAAGTATCATTTGCTGCTGGTGAATCAATGGTCAAGATCTCAAGATGGACTGCAAGGAG 840
Qy 1146 TACAACATGACAAAGACATTTGGAGAGTGGACAGTGGCACCAACCTTCGTTGCCCAAGAAA 1205
Db 841 TACAACATGACAAAGACATTTGGAGAGTGGACAGTGGCACCAACCTTCGTTGCCCAAGAAA 900
Qy 1206 GTCTTTGAAGCTGAGTCAATCCATCAAGGAGCTCTCCACGAGAGAGTTCCTGAT 1265
Db 901 GTATTTGAAGCTGCGCTCAAGTCCATCAAGGAGCTCTCCACGAGAGTTCCTGAT 960
Qy 1266 GTTTCTTGGCTAGGAGAGCAGTGGTGTGCTGGCAGCAGCACCCTTCGGAACATT 1325
Db 961 GGCTTTTGGCTAGGAGAGCAGTGGTGTGCTGGCAGCAGCAGCACCCTTCGGAACATT 1020
Qy 1326 TTCCCACTCATCTCACTTACCTTAATGGGTGAGGTATACCAACAGTCTCTTCCGATCACC 1385
Db 1021 TTCCCACTCATCTCACTTACCTTAATGGGTGAGGTATACCAACAGTCTCTTCCGATCACC 1080
Qy 1386 ATCTTTCCGAGCAATACCTCGCGCCAGTGGAGATGTGGCCAGTCCCAAGACGACTGT 1445
Db 1081 ATCTTTCCCTCAGCAATACCTTACGCGCGGTGGAGAGCTGGCCACGTCCTCAAGACGACTGT 1140
Qy 1446 TACAAGTTTCCCATCTCACAGTCATCCACGGGCACTGTATGGAGCTGTATATCATGGAG 1505
Db 1141 TACAAGTTTCCCTGTCTCACAGTCATCCACGGGCACTGTATGGAGCCGTATCATGGA 1200
Qy 1506 GGCTTTCTACGTTGTCTTTGATCGGGCCGAAACGAAATGGCTTTGCTGTACGGCTTC 1565
Db 1201 GGTTTCTATGCTGCTTCGATCGAGCCCGGAAAGCGAAATGGCTTTGCTGTACGGCTTC 1260
Qy 1566 CATGTGCACCATGATTCAGAGCGGAGCGGTGGAGGCTTTTGTCACTTTGGACATG 1625
Db 1261 CATGTGCACCATGATTCAGAGCGGCGGAGTGGAGGCTTCGTTTGTACGGGAGCATG 1320
Qy 1626 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCTTCATGACCATAGCCAT 1685
Db 1321 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCTTCATGACCATAGCCAT 1380
Qy 1686 GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCCTCATGCTGTACAGTGG 1745

Db 1381 GTCATGGCGGCATCTGCGCCCTTTCATGTTGCCACTCTGCCTCATGTATGTACAGTGG 1440
Qy 1746 CGTGTCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTCTG 1805
Db 1441 CGTGTCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTCTG 1500
Qy 1806 AAGTGA 1811
Db 1501 AAGTGA 1506
RESULT 14
US-09-548-372D-21
; Sequence 21, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-372D-21
Query Match 55.1%; Score 1293.8; DB 4; Length 1341;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 368 CACCCAGCAGCGCATCCGGCTGCCCCCTGCGCAGCGGCTGGGGGCGCCCCCTTGGGCT 427
Db 42 CACCCAGCAGCGCATCCGGCTGCCCCCTGCGCAGCGGCTGGGGGCGCCCCCTTGGGCT 101
Qy 428 GCGGCTGCCCCGGGAGACCGAGAGAGCCGCGGAGAGCCGCGGCGGAGGCGAGCTTTGT 487
Db 102 GCGGCTGCCCCGGGAGACCGAGAGAGCCGCGGAGAGCCGCGGCGGAGGCGAGCTTTGT 161
Qy 488 GGAGATGCTGGACAACTGAGGGGCAAGTGGGGCAGGCTACTACCTGGAGATGACCGT 547
Db 162 GGAGATGCTGGACAACTGAGGGGCAAGTGGGGCAGGCTACTACCTGGAGATGACCGT 221
Qy 548 GGCAGAGCCCCCGCAGACGCTCAACATCTCTGTTGATACAGGACGAGTAACTTTGCACT 607
Db 222 GGCAGAGCCCCCGCAGACGCTCAACATCTCTGTTGATACAGGACGAGTAACTTTGCACT 281
Qy 608 GGTGCTGCCCCCGCAGACGCTTCTGATCGTACTACAGAGGAGCTGTCCAGACATA 667
Db 282 GGTGCTGCCCCCGCAGACGCTTCTGATCGTACTACAGAGGAGCTGTCCAGACATA 341
Qy 668 CCGGAGCTCCGGAAGGCTGTATGTGCTTACACCCAGGCAAGTGGGAAGGGAGCT 727
Db 342 CCGGAGCTCCGGAAGGCTGTATGTGCTTACACCCAGGCAAGTGGGAAGGGAGCT 401
Qy 728 GGCACACGACCTGTTAAGCATCCCTTACCTGCCCCAACGTCACCTGCTGCTGCAACATTGC 787
Db 402 GGCACACGACCTGTTAAGCATCCCTTACCTGCCCCAACGTCACCTGCTGCTGCAACATTGC 461
Qy 788 TGCCATCACTGAATCAGACAAAGTTCCTTCAACAGGCTTCCAACTGGGAAGGCACTCTGGG 847
Db 462 TGCCATCACTGAATCAGACAAAGTTCCTTCAACAGGCTTCCAACTGGGAAGGCACTCTGGG 521

QY 848 GCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTCTTTTGACTCTCT 907
D 522 GCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTCTTTTGACTCTCT 581
QY 908 GGTAAAGCAGACCCAGCTTCCCAACCTCTCTCCCTGACGCTTTGTGGTGTGGCTTCCC 967
D 582 GGTAAAGCAGACCCAGCTTCCCAACCTCTCTCCCTGACGCTTTGTGGTGTGGCTTCCC 641
QY 968 CCTCAACGAGTCTGAAGTCTGGCTCTCTCGGAGGAGCATGATCATTTGGAGGTATCGA 1027
D 642 CCTCAACGAGTCTGAAGTCTGGCTCTCTCGGAGGAGCATGATCATTTGGAGGTATCGA 701
QY 1028 CCACTCGCTGACACAGCAGTCTCTGGTATACACCCATCCGCGGAGTGGTATTTATGA 1087
D 702 CCACTCGCTGACACAGCAGTCTCTGGTATACACCCATCCGCGGAGTGGTATTTATGA 761
QY 1088 GGTGATCATTTGGGCTGGAGTCAATGGACAGGATCTGAATGAGCTGCAAGGAGTA 1147
D 762 GGTGATCATTTGGGCTGGAGTCAATGGACAGGATCTGAATGAGCTGCAAGGAGTA 821
QY 1148 CAATATGACAAGAGCATTTGTGGACAGTGGCACCACCAACCTTCTCTTTGCCCAAGAAAGT 1207
D 822 CAATATGACAAGAGCATTTGTGGACAGTGGCACCACCAACCTTCTCTTTGCCCAAGAAAGT 881
QY 1208 GTTTGAAGTCCAGTCAATCCATCAAGCAGCCTCTCCACGAGAGTTCCTGTATGG 1267
D 882 GTTTGAAGTCCAGTCAATCCATCAAGCAGCCTCTCCACGAGAGTTCCTGTATGG 941
QY 1268 TTTCTGCTAGGAGCAGCTGGTGTGGCAGGAGGACACCCCTTGGACATTTT 1327
D 942 TTTCTGCTAGGAGCAGCTGGTGTGGCAGGAGGACACCCCTTGGACATTTT 1001
QY 1328 CCCAGTCTACCTACCTACCTAATGGGTGAGGTATACCAACGAGTTCCTCCGATCACCAT 1387
D 1002 CCCAGTCTACCTACCTAATGGGTGAGGTATACCAACGAGTTCCTCCGATCACCAT 1061
QY 1388 CTTTCCGAGCAATACCTGCGGCCAGTGGAAAGATGGCCACGCTTCCCAAGACGACTGTTA 1447
D 1062 CTTTCCGAGCAATACCTGCGGCCAGTGGAAAGATGGCCACGCTTCCCAAGACGACTGTTA 1121
QY 1448 CAAGTTTGGCATCTACAGTCTACACGGGCTGTTATGGAGCTGTTATCATGGAGG 1507
D 1122 CAAGTTTGGCATCTACAGTCTACACGGGCTGTTATGGAGCTGTTATCATGGAGG 1181
QY 1508 CTTTACGTTGCTTTGATCGGGCCGAAACGAAATGGCTTTGCTGTCAGCGCTTGCCA 1567
D 1182 CTTTACGTTGCTTTGATCGGGCCGAAACGAAATGGCTTTGCTGTCAGCGCTTGCCA 1241
QY 1568 TGTGACGATGAGTTCAGGAGGCGGAGCGGTGGAGGCCCTTTGTACCTTGGACATGGA 1627
D 1242 TGTGACGATGAGTTCAGGAGGCGGAGCGGTGGAGGCCCTTTGTACCTTGGACATGGA 1301
QY 1628 AGACTGTGGCTACACATTTCCACACACATGAGTCA 1664
D 1302 AGACTGTGGCTACACATTTCCACACACATGAGTCA 1338

RESULT 15
US-09-548-367D-21
: Sequence 21, Application US/09548367D
: Patent No. 6440698
: GENERAL INFORMATION:
: APPLICANT: GURNEY ET AL.
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
: FILE REFERENCE: 29915/6280H
: CURRENT APPLICATION NUMBER: US/09/548,367D
: CURRENT FILING DATE: 2000-04-12
: PRIOR APPLICATION NUMBER: US 60/155,493
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: US 09/404,133
: PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: PCT/US99/20881
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: US 60/101,594
: PRIOR FILING DATE: 1998-09-24
: NUMBER OF SEQ ID NOS: 73
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 21
: LENGTH: 1341
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-548-367D-21

Query Match 55.1%; Score 1293.8; DB 4; Length 1341;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 368 CACCAGCAGCAGCATCCGGCTGCCCTTGCAGAGGGCTGGGGGCGCCCTGGGGCT 427
D 42 CACCAGCAGCAGCATCCGGCTGCCCTTGCAGAGGGCTGGGGGCGCCCTGGGGCT 101
QY 428 GCGGTGCTCCCGGAGACCGACGAAGAGCCGAGGAGCCCGCGGAGGGGAGCTTTGT 487
D 102 GCGGTGCTCCCGGAGACCGACGAAGAGCCGAGGAGCCCGCGGAGGGGAGCTTTGT 161
QY 488 GGAGATGTGTGAACACCTGAGGGGCAAGTCGGGCGAGGGCTTACTACGTGGAGATGACCGT 547
D 162 GGAGATGTGTGAACACCTGAGGGGCAAGTCGGGCGAGGGCTTACTACGTGGAGATGACCGT 221
QY 548 GGGCAGCCCCCGCAGACGCTCAACATCTCTGGTGGATACAGGCAGCAGTAACCTTTCAGT 607
D 222 GGGCAGCCCCCGCAGACGCTCAACATCTCTGGTGGATACAGGCAGCAGTAACCTTTCAGT 281
QY 608 GGGTGTCTCCCGCCCGCCTTCTGTCATCGCTACTACAGAGGAGCTGTCCAGCACATA 667
D 282 GGGTGTCTCCCGCCCGCCTTCTGTCATCGCTACTACAGAGGAGCTGTCCAGCACATA 341
QY 668 CCGGACCTCCGGAAGGTGTGTATGTGCTTACACCCAGGGCAAGTGGGAAGGGAGCT 727
D 342 CCGGACCTCCGGAAGGTGTGTATGTGCTTACACCCAGGGCAAGTGGGAAGGGAGCT 401
QY 728 GGGCAGCCAGCTTGTAAAGCATCCCGCATGCCCAAGCTACTGTGCGTGGCAACATTCG 787
D 402 GGGCAGCCAGCTTGTAAAGCATCCCGCATGCCCAAGCTACTGTGCGTGGCAACATTCG 461
QY 788 TGCCATCACTGAATCAGACAAGTTCTTCATCAAGGGCTCCAACTGGGAAGGCATCTTGG 847
D 462 TGCCATCACTGAATCAGACAAGTTCTTCATCAAGGGCTCCAACTGGGAAGGCATCTTGG 521
QY 848 GCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCTCT 907
D 522 GCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCTCT 581
QY 908 GGTAAAGCAGACCCAGCTTCCCAACCTCTCTCCCTGACGCTTTGTGGTGTGGCTTCCC 967
D 582 GGTAAAGCAGACCCAGCTTCCCAACCTCTCTCCCTGACGCTTTGTGGTGTGGCTTCCC 641
QY 968 CCTCAACGAGTCTGAAGTCTGGCTCTCTCGGAGGAGCATGATCATTTGGAGGTATCGA 1027
D 642 CCTCAACGAGTCTGAAGTCTGGCTCTCTCGGAGGAGCATGATCATTTGGAGGTATCGA 701
QY 1028 CCACTCGCTGACACAGGAGTCTCTGGTATACACCCATCCGCGGAGTGGTATTTATGA 1087
D 702 CCACTCGCTGACACAGGAGTCTCTGGTATACACCCATCCGCGGAGTGGTATTTATGA 761
QY 1088 GGTGATCATTTGGGCTGGAGTCAATGGACAGGATCTGAATGAGCTGCAAGGAGTA 1147
D 762 GGTGATCATTTGGGCTGGAGTCAATGGACAGGATCTGAATGAGCTGCAAGGAGTA 821
QY 1148 CAATATGACAAGAGCATTTGTGGACAGTGGCACCACCAACCTTCTCTTTGCCCAAGAAAGT 1207
D 822 CAATATGACAAGAGCATTTGTGGACAGTGGCACCACCAACCTTCTCTTTGCCCAAGAAAGT 881
QY 1208 GTTTGAAGTCCAGTCAATCCATCAAGCAGCCTCTCCACGAGAGTTCCTGTATGG 1267

```
|||||
Db 882 GTTTGAAGCTGCAGTCAAAATCCATCAAGGCAGCCTCTCCACGGAGAAGTTCCCTGATGG 941
QY 1268 TTTCTGGCTAGGAGAGCAGTGTGTGCTGGCAGCAGGACACCCCTTTGGAAACATTTT 1327
Db 942 TTTCTGGCTAGGAGAGCAGTGTGTGCTGGCAAGCAGCAGCACCACCCTTTGGAAACATTTT 1001
QY 1328 CCCAGTCATCTCACTCTACCTAATGGGTGAGGTTTACCAACAGTCCTTCCGCATCAGCAT 1387
Db 1002 CCCAGTCATCTCACTCTACCTAATGGGTGAGGTTTACCAACAGTCCTTCCGCATCAGCAT 1061
QY 1388 CCTTCCGACAGCAATACCTGCGGCCAGTGGGAAGATGTGGCCACGTCCTCCCAAGACGACTGTTA 1447
Db 1062 CCTTCCGACAGCAATACCTGCGGCCAGTGGGAAGATGTGGCCACGTCCTCCCAAGACGACTGTTA 1121
QY 1448 CAAGTTTGGCCATCTCACAGTCATCCACGGGCAGTGTATGGGAGCTGTTATCATGGAGGG 1507
Db 1122 CAAGTTTGGCCATCTCACAGTCATCCACGGGCAGTGTATGGGAGCTGTTATCATGGAGGG 1181
QY 1508 CTCTACGTTGTCTTTTCATCGGGCCGAAACGAATTTGGCTTTGCTGTCAGCGCTTGCCA 1567
Db 1182 CTCTACGTTGTCTTTTCATCGGGCCGAAACGAATTTGGCTTTGCTGTCAGCGCTTGCCA 1241
QY 1568 TGTGCACGATGAGTTCAGGACGGCGGTGGAAAGGCCCTTTTGTCAACCTTGGACATGGA 1627
Db 1242 TGTGCACGATGAGTTCAGGACGGCGGTGGAAAGGCCCTTTTGTCAACCTTGGACATGGA 1301
QY 1628 AGACTGTGGCTACAAACATTCACAGACAGATGAGTCA 1664
Db 1302 AGACTGTGGCTACAAACATTCACAGACAGATGAGTCA 1338
```

Search completed: March 2, 2003, 06:15:53
Job time : 79 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2003, 13:37:20 ; Search time 324.5 Seconds
(without alignments)
16294.891 Million cell updates/sec

Title: US-09-723-722A-42
Perfect score: 2348
Sequence: 1 ccatcgccgcccctcacagc.....caagattgctcttggaatt 2348

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_101002:*
1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*			
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*			
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*			
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*			
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*			
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*			
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*			
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*			
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*			
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*			
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*			
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*			
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*			
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*			
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*			
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*			
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*			
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*			
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*			
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*			
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*			
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*			
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*			
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length DB	ID	Description
1	2347.6	100.0	2348	21	AAA59551
2	2336.4	99.5	16080	21	AAA59553
3	2181.2	92.9	5757	24	ABL39774
4	2039.8	86.9	2070	21	AAAL5662
5	2039.8	86.9	2070	22	AAD17865
6	2039.8	86.9	2070	22	AAD13021
7	2039.8	86.9	2070	22	AAS11517
8	2039.8	86.9	2070	22	AAS11702
9	2039.8	86.9	2070	22	AAD06739

10	2039.8	86.9	2070	24	ABL52457	Human Asp-2(a) nuc
11	2039.8	86.9	2070	24	ABL49914	Human aspartyl pro
12	2039.8	86.2	2541	19	AAV41696	Nucleotide sequenc
13	2004	85.3	3252	22	AAF31848	Human memapsin 2 c
14	2004	85.3	3252	22	AAF28101	CDNA encoding huma
15	2004	85.3	3252	24	ABK8641	Memapsin 2 DNA, H
16	1881	80.1	1977	21	AAAL5663	Human aspartyl pro
17	1881	80.1	1977	21	AAD17866	Human aspartyl pro
18	1881	80.1	1977	22	AAD13022	Human aspartyl pro
19	1881	80.1	1977	22	AAS11518	Human cDNA encodin
20	1881	80.1	1977	22	AAS11703	DNA encoding human
21	1881	80.1	1977	22	AAD06740	Human aspartyl pro
22	1881	80.1	1977	24	ABL52458	Human Asp-2(b) nuc
23	1881	80.1	1977	24	ABL49915	Human aspartyl pro
24	1835.4	78.2	2370	19	AAV41697	Partial nucleotide
25	1792	76.3	1379	22	AAK94824	Human full-length
26	1648	70.2	1747	20	AAK97602	Extended human sec
27	1550	66.0	2043	21	AAAL5664	Murine aspartyl pr
28	1550	66.0	2043	22	AAD17867	Murine aspartyl pr
29	1550	66.0	2043	22	AAD13023	Murine aspartyl pr
30	1550	66.0	2043	22	AAS11519	Mouse cDNA encodin
31	1550	66.0	2043	22	AAS11704	DNA encoding mouse
32	1550	66.0	2043	22	AAD06741	Murine aspartyl pr
33	1550	66.0	2043	24	ABL52459	Mouse Asp-2(a) nuc
34	1511.2	64.4	2158	24	ABK63758	Rat sequence diffe
35	1504.4	64.1	2907	23	AAK82237	DNA encoding novel
36	1503	64.0	1503	21	AAK28278	Human cDNA encodin
37	1503	64.0	1503	21	AAA59550	DNA encoding a hum
38	1503	64.0	1527	24	ABA02406	FLAG-tagged human
39	1476.2	62.9	1911	22	AAD09485	Human aspartyl pro
40	1423	60.6	2914	23	AAK37398	DNA encoding novel
41	1355.8	57.7	1362	21	AAAL5668	Modified human asp
42	1355.8	57.7	1362	22	AAD17878	Human-Asp 2(a) pro
43	1355.8	57.7	1362	22	AAD13034	Human-Asp2(a) delt
44	1355.8	57.7	1362	22	AAS11530	Human cDNA encodin
45	1355.8	57.7	1362	22	AAS11715	DNA encoding human

ALIGNMENTS

RESULT 1			
AAA59551			
ID	AAA59551	standard; DNA; 2348 BP.	
XX			
AC	AAA59551;		
XX			
DT	14-NOV-2000	(first entry)	
XX			
DE	DNA encoding a human beta-secretase enzyme.		
XX			
KW	Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;		
KW	amyloid plaque component; Alzheimer's disease; amyloidogenic disease;		
KW	inhibitor; ss.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	306..1811	
FT		/*tag= a	
FT		/product= "beta-secretase"	
XX			
PN	WO200047618-A2.		
XX			
PD	17-AUG-2000.		
XX			
PF	10-FEB-2000; 2000WO-US03819.		
XX			
PR	10-FEB-1999; 99US-0119571.		
PR	15-JUN-1999; 99US-0139172.		
XX			
PA	(ELAN-) ELAN PHARM INC.		
XX			

PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX WPI: 2000-533011/48.
DR P-PSDB: AAB07896.
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX Disclosure; Fig 1B; 121pp; English.
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence encodes a human beta-secretase enzyme.
XX
SQ Sequence 2348 BP: 489 A; 713 C; 661 G; 484 T; 1 other;

Query Match 100.0%; Score 2347.6; DB 21; Length 2348;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 2348; Conservative 0; Indels 0;

QY 1 CCATGCGGGCCCTCACAGCCCGCGGGAGCCCGAGCCGCTGCCAGGCTGGCCGCG 60
DB 1 CCATGCGGGCCCTCACAGCCCGCGGGAGCCCGAGCCGCTGCCAGGCTGGCCGCG 60

QY 61 CSFTGCCGATGAGGGGCTCGGATGCCAGCTCTCCCTGCTCCCGTGTCTGCGGAT 120
DB 61 CSFTGCCGATGAGGGGCTCGGATGCCAGCTCTCCCTGCTCCCGTGTCTGCGGAT 120

QY 121 CTCCCTGACCGCTTCACAGCCCGGACCCGGGGGTGGCCAGGCGCTTCAGGCGCT 180
DB 121 CTCCCTGACCGCTTCACAGCCCGGACCCGGGGGTGGCCAGGCGCTTCAGGCGCT 180

QY 181 GCGTCTGTATGCCCCCAAGCTCCCTCTCTGAGAGCCACAGCACCCAGACTTGG 240
DB 181 GCGTCTGTATGCCCCCAAGCTCCCTCTCTGAGAGCCACAGCACCCAGACTTGG 240

QY 241 GGCAGCGCCGAGGACGACGTGGGCCAGTGGAGCCAGAGGCGCCGAGGCGGGGC 300
DB 241 GGCAGCGCCGAGGACGACGTGGGCCAGTGGAGCCAGAGGCGCCGAGGCGGGGC 300

QY 301 CCACCATGGCCCAAGCCCTGCCCTGGCTCTCTGATGGATGGGGCGGGAGTGTGCGTG 360
DB 301 CCACCATGGCCCAAGCCCTGCCCTGGCTCTCTGATGGATGGGGCGGGAGTGTGCGTG 360

QY 361 CCCAGCCACCCAGCAGCGCATCCGCTGCCCTGGCAGCGCTGGGGCGGCCCGCC 420
DB 361 CCCAGCCACCCAGCAGCGCATCCGCTGCCCTGGCAGCGCTGGGGCGGCCCGCC 420

QY 421 TGGGGCTGCGGCTGCCCGGGAGACCGAGCAGAGCCCGAGGCGCCGCGGAGGGGCA 480
DB 421 TGGGGCTGCGGCTGCCCGGGAGACCGAGCAGAGCCCGAGGCGCCGCGGAGGGGCA 480

QY 481 GCATTGTGGAGATGGTGACAACTTGAAGGGGCAAGTGGGGGAGGGGTACTACGTGGAGA 540
DB 481 GCATTGTGGAGATGGTGACAACTTGAAGGGGCAAGTGGGGGAGGGGTACTACGTGGAGA 540

QY 541 TGACCGTGGGCGCCCGCCGAGCGCTCAACATCTGTGTGATACAGGAGCAGTAAGT 600
DB 541 TGACCGTGGGCGCCCGCCGAGCGCTCAACATCTGTGTGATACAGGAGCAGTAAGT 600

QY 601 TTGAGTGGGTGCTGCCCGCCCGCCCGCTTCTGATCGCTACTACAGAGCGAGCTGTCCA 660
DB 601 TTGAGTGGGTGCTGCCCGCCCGCCCGCTTCTGATCGCTACTACAGAGCGAGCTGTCCA 660

QY 661 GCACATACCGGGACCTCCCGAAGGGTGTGTATGTGCCCTACACCCAGGSCAAGTGGGAAG 720
DB 661 GCACATACCGGGACCTCCCGAAGGGTGTGTATGTGCCCTACACCCAGGSCAAGTGGGAAG 720

QY 721 GGGAGCTGGGACCGGACCTGGTAAGCATCCCCATGGCCCCCAACGCTCACTGCTGCGTGGCA 780
DB 721 GGGAGCTGGGACCGGACCTGGTAAGCATCCCCATGGCCCCCAACGCTCACTGCTGCGTGGCA 780

QY 781 ACATTGCTGCCATCACTGAATCAGACAAGTCTTCTCATCAAGGCTCCAACTGGGAAGGCA 840
DB 781 ACATTGCTGCCATCACTGAATCAGACAAGTCTTCTCATCAAGGCTCCAACTGGGAAGGCA 840

QY 841 TCCTGGGGTGGCCTATGTGATGATGCCAGGCGCTGAGACTCTCCCTGGAGCCTTCTTTG 900
DB 841 TCCTGGGGTGGCCTATGTGATGATGCCAGGCGCTGAGACTCTCCCTGGAGCCTTCTTTG 900

QY 901 ACTCTCTGTTAAAGCAGACCCAGCTTCCCAACCTTCTCTCCCTGACGCTTTGTGGTGGCTG 960
DB 901 ACTCTCTGTTAAAGCAGACCCAGCTTCCCAACCTTCTCTCCCTGACGCTTTGTGGTGGCTG 960

QY 961 GCTTCCCTCTCAACCAAGTCTGAAGTGTGGCTCTGTCGGAGGAGCATGATCATTTGGAG 1020
DB 961 GCTTCCCTCTCAACCAAGTCTGAAGTGTGGCTCTGTCGGAGGAGCATGATCATTTGGAG 1020

QY 1021 GTATCGACCACTCGGTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGGT 1080
DB 1021 GTATCGACCACTCGGTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGGT 1080

QY 1081 ATTATGAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCA 1140
DB 1081 ATTATGAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCA 1140

QY 1141 AGGAGTACAACATATGACAAGAGCATTTGGACAGTGGCACCACCAACCTTCGTTTCCCA 1200
DB 1141 AGGAGTACAACATATGACAAGAGCATTTGGACAGTGGCACCACCAACCTTCGTTTCCCA 1200

QY 1201 AGAAAGTGTTCGAAGCTGCAGTCAAAATCCATCAAGAGCAGCTCTCCACGGAGAGTTC 1260
DB 1201 AGAAAGTGTTCGAAGCTGCAGTCAAAATCCATCAAGAGCAGCTCTCCACGGAGAGTTC 1260

QY 1261 CTGATGGTTCCTGGCTAGGAGAGCAGTGGTGTGCTGGCAGCAGGACCCCTTGA 1320
DB 1261 CTGATGGTTCCTGGCTAGGAGAGCAGTGGTGTGCTGGCAGCAGGACCCCTTGA 1320

QY 1321 ACATTTTCCAGTCACTCTACCTAATGGTGGAGTTCACCAACAGCTCTTCCCGCA 1380
DB 1321 ACATTTTCCAGTCACTCTACCTAATGGTGGAGTTCACCAACAGCTCTTCCCGCA 1380

QY 1381 TCACCATCTTCGGCAGCAATACCTGGGCGCAGTGGAGAGTGTGGCCACGTCCTCCAGAG 1440
DB 1381 TCACCATCTTCGGCAGCAATACCTGGGCGCAGTGGAGAGTGTGGCCACGTCCTCCAGAG 1440

QY 1441 ACTGTTACAAGTTTGGCATCTCAGTCATCCAGGCGCAGTGTATGGAGCTGTATCA 1500
DB 1441 ACTGTTACAAGTTTGGCATCTCAGTCATCCAGGCGCAGTGTATGGAGCTGTATCA 1500

QY 1501 TGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCCAAAACAAATTTGGCTGTGTCAGCG 1560
DB 1501 TGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCCAAAACAAATTTGGCTGTGTCAGCG 1560

QY 1561 CTTGGCATGTGACAGATGAGTTCAGGAGCGGCGGTGGAGGGCCCTTTTGTACCTTGG 1620
DB 1561 CTTGGCATGTGACAGATGAGTTCAGGAGCGGCGGTGGAGGGCCCTTTTGTACCTTGG 1620

QY 1621 ACATGGAAGACTGTGCTACACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG 1680
DB 1621 ACATGGAAGACTGTGCTACACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG 1680

QY 1681 CCTATGTCATGGCTGCCCTCTCTTATGCTGCCACTCTCTGCTCATGCTGTGTC 1740
DB 1681 CCTATGTCATGGCTGCCCTCTCTTATGCTGCCACTCTCTGCTCATGCTGTGTC 1740

QY 1741 AGTGGCGCTGCCCTGCCCTGCCCGCAGCAGCATGATGACTTTGCTGATGACATCTCC 1800

Db 1741 AGTGGCGCTGCTCCGCTGCTGCGCCAGCATGATGATTTGCTGATGACATCTCCC 1800
Qy 1801 TGCTGAAGTCAGAGGCCATGGCCAGCAAGATAGAGATTTCCCTGGACCACACCTCCGCTG 1860
Db 1801 TGCTGAAGTCAGAGGCCATGGCCAGCAAGATAGAGATTTCCCTGGACCACACCTCCGCTG 1860
Qy 1861 GTTCACATTTGGTCACAAAGTAGGACACAGATGGCACCTGTGGCCAGAGACCTCAGGAC 1920
Db 1861 GTTCACATTTGGTCACAAAGTAGGACACAGATGGCACCTGTGGCCAGAGACCTCAGGAC 1920
Qy 1921 CTTCCCCACCCACCAATGCTTCTGCTTGTATGAGAGAAAGGCTGGCAAGGTGGGT 1980
Db 1921 CTTCCCCACCCACCAATGCTTCTGCTTGTATGAGAGAAAGGCTGGCAAGGTGGGT 1980
Qy 1981 TCCAGGACATGCTACCTCTAGGAAACAGAAAGAGAAAGCAAGCACTCTGCTGGCGGA 2040
Db 1981 TCCAGGACATGCTACCTCTAGGAAACAGAAAGAGAAAGCAAGCACTCTGCTGGCGGA 2040
Qy 2041 ATACTCTTGGTCACTCAAAATTTAAAGTCGGGAAATTTCTGCTTGAACCTTCAGCCCTG 2100
Db 2041 ATACTCTTGGTCACTCAAAATTTAAAGTCGGGAAATTTCTGCTTGAACCTTCAGCCCTG 2100
Qy 2101 AACCTTTGTCACATTTCTTAAATTTCTCAACCCAAAGATATCTTTCTTTAGTTT 2160
Db 2101 AACCTTTGTCACATTTCTTAAATTTCTCAACCCAAAGATATCTTTCTTTAGTTT 2160
Qy 2161 CAGAAGTACTGGCATACAGCAGGTACCTTGCGTGTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2161 CAGAAGTACTGGCATACAGCAGGTACCTTGCGTGTGCTGCTGCTGCTGCTGCTGCTG 2220
Qy 2221 AGAAGACACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG 2280
Db 2221 AGAAGACACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG 2280
Qy 2281 CTATTTGCTTTAGACAGGAGCTGTATAAACAAAGCCTTAACATTTGTTGCAAAAGATTGCC 2340
Db 2281 CTATTTGCTTTAGACAGGAGCTGTATAAACAAAGCCTTAACATTTGTTGCAAAAGATTGCC 2340
Qy 2341 CTTGAATT 2348
Db 2341 CTTGAATT 2348

RESULT 2
AAA59553
ID AAA59553 standard; DNA; 16080 BP.

XX AC AAA59553;

XX DT 14-NOV-2000 (first entry)

XX DE DNA clone pCEK Cl.27 encoding a human beta-secretase enzyme.

XX KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW Inhibitor; ss.

XX OS Homo sapiens.

XX PN W0200047618-A2.

XX XX 17-AUG-2000.

XX XX 10-FEB-2000; 2000WO-US03819.

XX XX 10-FEB-1999; 99US-0119571.

XX XX 15-JUN-1999; 99US-0139172.

XX PA (ELAN-) ELAN PHARM INC.

XX PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;

XX PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;

XX WPI; 2000-533011/48.
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
XX which can be used for the treatment of amyloidogenic diseases e.g.
XX Alzheimer's disease -
XX
XX Disclosure: Fig 13A-E; 12lpp; English.
XX
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
XX beta-amyloid precursor protein to produce beta-amyloid peptide. This
XX enzyme is therefore implicated in the production of amyloid plaque
XX components which accumulate in the brains of individuals afflicted with
XX Alzheimer's disease. Inhibitors of beta-secretase are administered to
XX a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
XX disease-like pathology to test if they maintain or improve cognitive
XX ability or reduce the plaque burden. The compounds are used for the
XX treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
XX present sequence encodes a human beta-secretase enzyme.
XX
XX Sequence 16080 BP; 3627 A; 4556 C; 3962 G; 3913 T; 22 other;

Query Match 99.5%; Score 2336.4; DB 21; Length 16080;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2340; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Qy 1 CCATGCGCGCCCTCACAGCCCGCGGAGCCGCGCTGCCAGGCTGGCGCGC 60
Db 1370 CCATGCGCGCCCTCCAGCCCCCGCGGAGCCGCGCTGCCAGGCTGGCGCGC 1429
Qy 61 CSGTGCGGATGTAGCGGCTCCGGATCCAGCCTCTCCCTGCTCCCTGCTCTGCGGAT 120
Db 1430 CCGTGGCGATGTAGCGGCTCCGGATCCAGCCTCTCCCTGCTCTGCGGAT 1489
Qy 121 CTCCTCTGACCGCTCTCACAGCCCGGAGCCGCGGCTGCGCCAGGCGCTTCAGGCGCT 180
Db 1490 CTCCTCTGACCGCTCTCACAGCCCGGAGCCGCGGCTGCGCCAGGCGCTTCAGGCGCT 1549
Qy 181 GCGCTCTGATGCCCCCAAGCTCTCTCTGAGAAGCCACAGCACCACACACTTGG 240
Db 1550 GCGCTCTGATGCCCCCAAGCTCTCTCTGAGAAGCCACAGCACCACACTTGG 1609
Qy 241 GGCAGGCGCCAGGAGCGAGTGGGCGAGTGCAGGCCAGAGGCCCGCCAGGCGCGGC 300
Db 1610 GGCAGGCGCCAGGAGCGAGTGGGCGAGTGCAGGCCAGAGGCCCGCCAGGCGCGGC 1669
Qy 301 CCACATGGCCCAAGCGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 1670 CCACATGGCCCAAGCGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729
Qy 361 CCCAGGCGCCAGGAGCGAGTGGGCGAGTGCAGGCCAGAGGCCCGCCAGGCGCGGC 420
Db 1730 CCCAGGCGCCAGGAGCGAGTGGGCGAGTGCAGGCCAGAGGCCCGCCAGGCGCGGC 1789
Qy 421 TGGGGCTGCGGCTGCCCGGAGACCGAGCAAGAGCCCGAGAGGCCCGCCGAGGGGCA 480
Db 1790 TGGGGCTGCGGCTGCCCGGAGACCGAGCAAGAGCCCGAGAGGCCCGCCGAGGGGCA 1849
Qy 481 GCTTTTGGAGATGGTGACAACTGAGGGGCAAGTGGGGCGGCGCTACTACGTGGAGA 540
Db 1850 GCTTTTGGAGATGGTGACAACTGAGGGGCAAGTGGGGCGGCGCTACTACGTGGAGA 1909
Qy 541 TGACCGTGGGCGAGCCCCCGCGAGAGCTCAACATCTCTGCTGATACAGGAGCAGTAAC 600
Db 1910 TGACCGTGGGCGAGCCCCCGCGAGAGCTCAACATCTCTGCTGATACAGGAGCAGTAAC 1969
Qy 601 TTGAGTGGGTGCTGCCCGCCACCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 1970 TTGAGTGGGTGCTGCCCGCCACCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2029
Qy 661 GCACATACCGGAGCTCCGGAGAGGTGTGTATGTGCCCTTACACCCAGGCGAGTGGGAG 720
Db 2030 GCACATACCGGAGCTCCGGAGAGGTGTGTATGTGCCCTTACACCCAGGCGAGTGGGAG 2089

Qy	721	GGGAGCTGGGCACCGACCTGGTAAGCATCCCCATGGCCCCCAACGTCACCTGTGCGGTGCCA	780
Db	2090	GGGAGCTGGGCACCGACCTGGTAAGCATCCCCATGGCCCCCAACGTCACCTGTGCGGTGCCA	2149
Qy	781	ACATTTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCA	840
Db	2150	ACATTTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCA	2209
Qy	841	TCCTGGGCTGGCCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGGCCCTTCTTTG	900
Db	2210	TCCTGGGCTGGCCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGGCCCTTCTTTG	2269
Qy	901	ACTCTCTGTTAAGCAGACCCACGTTTCCAACCTCTTCTCCCTGACGCTTTGTGTGCTGTG	960
Db	2270	ACTCTCTGTTAAGCAGACCCACGTTTCCAACCTCTTCTCCCTGACGCTTTGTGTGCTGTG	2329
Qy	961	CGTTCGCCCTCAACCAGTCTGAAGTCTGCTGCGGAGGGAGCATGATCATTTGGAG	1020
Db	2330	CGTTCGCCCTCAACCAGTCTGAAGTCTGCTGCGGAGGGAGCATGATCATTTGGAG	2389
Qy	1021	GTATCGACCACTCGCTGTACACAGGCAGTCTCTCTGGTATACACCCATCCGGCGGAGTGGT	1080
Db	2390	GTATCGACCACTCGCTGTACACAGGCAGTCTCTCTGGTATACACCCATCCGGCGGAGTGGT	2449
Qy	1081	ATTATGAGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCA	1140
Db	2450	ATTATGAGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCA	2509
Qy	1141	AGGAGTACAACATATCACAAAGACCAATTGGACAGTGGCACCAACCAACCTTCGTTTGC	1200
Db	2510	AGGAGTACAACATATGACAAGACCAATTGGACAGTGGCACCAACCAACCTTCGTTTGC	2569
Qy	1201	AGAAAGTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCCTCTCCAGGGAGAACTTC	1260
Db	2570	AGAAAGTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCCTCTCCAGGGAGAACTTC	2629
Qy	1261	CTGATGGTTTCTGGCTAGGAGACGACTGGTGTGCTGGCAAGCAGGACCAACCCCTTGG	1320
Db	2630	CTGATGGTTTCTGGCTAGGAGACGACTGGTGTGCTGGCAAGCAGGACCAACCCCTTGG	2689
Qy	1321	ACATTTTCCCAGTCTATCTCACCTCTACCTAATGGGTGAGGTTACCAACCAAGTCTTCCG	1380
Db	2690	ACATTTTCCCAGTCTATCTCACCTCTACCTAATGGGTGAGGTTACCAACCAAGTCTTCCG	2749
Qy	1381	TCACCATCTTCCGAGCAAAATACCTGCGGCCAGTGAAGATGTGGCCACGCTCCCAAGACG	1440
Db	2750	TCACCATCTTCCGAGCAAAATACCTGCGGCCAGTGAAGATGTGGCCACGCTCCCAAGACG	2809
Qy	1441	ACTGTTACAAGTTTGCCTATCTCACAGTCAATCCACGGGCACTGTTATGGAGCTGTTATCA	1500
Db	2810	ACTGTTACAAGTTTGCCTATCTCACAGTCAATCCACGGGCACTGTTATGGAGCTGTTATCA	2869
Qy	1501	TGGAGGGCTTCTACGTTGTCTTTGATGCGGGCCGAAACGAATTTGGCTTTGTGTCTCAGCG	1560
Db	2870	TGGAGGGCTTCTACGTTGTCTTTGATGCGGGCCGAAACGAATTTGGCTTTGTGTCTCAGCG	2929
Qy	1561	CTTGCCATGTGCACGATGAGTTTCAGACGGCAGCGGTGGAAGGCCCTTTTGTACACTTGG	1620
Db	2930	CTTGCCATGTGCACGATGAGTTTCAGACGGCAGCGGTGGAAGGCCCTTTTGTACACTTGG	2989
Qy	1621	ACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG	1680
Db	2990	ACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG	3049
Qy	1681	CCTATGTCATGGCTGCCATCTCGGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGC	1740
Db	3050	CCTATGTCATGGCTGCCATCTCGGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGC	3109
Qy	1741	AGTGGCGTCCCTCGCTGCCTTGGCCAGCAGATGATGACTTTTGTGTGATGACATCTCCC	1800
Db	3110	AGTGGCGTCCCTCGCTGCCTTGGCCAGCAGATGATGACTTTTGTGTGATGACATCTCCC	3169

QY	1801	TGCTGAAGTGGAGGCCCATCGGCAGAAGATAGAGATTCCCTCTGACACACACTTCGTG	1861
Db	3170	TGCTGAAGTGGAGGCCCATCGGCAGAAGATAGAGATTCCCTTGACCACACACTTCGCTC	3229
QY	1861	GTTTCACTTTGGTTCACAAGTAGAGACACAGATGGCACCTGTGGCCAGACACCTCAGGAC	1920
Db	3230	GTTTCACTTTGGTTCACAAGTAGAGACACAGATGGCACCTGTGGCCAGACACCTCAGGAC	3289
QY	1921	CCTCCCCACCCACCAAATGCCCTTCGCTTGGATGGAGAAGAAAGAGCTGCCAAGTGGGT	1980
Db	3290	CCTCCCCACCCACCAAATGCCCTTCGCTTGGATGGAGAAGAAAGAGCTGCCAAGTGGGT	3349
QY	1981	TCCAGGGACTGTACCTGTAGGAAACAGAAAAGAGAAGAACGACTCTGCTGGCGGGA	2040
Db	3350	TCCAGGGACTGTACCTGTAGGAAACAGAAAAGAGAAGAACGACTCTGCTGGCGGGA	3409
QY	2041	ATACTCTTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTTGAACACTTCAGCCCTG	2100
Db	3410	ATACTCTTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTTGAACACTTCAGCCCTG	3469
QY	2101	AACCTTTGTCCACCACTTCCTTTAAATTTCCCAACCCAAAAGTATTCTCTTTCTTAGTTTT	2160
Db	3470	AACCTTTGTCCACCACTTCCTTTAAATTTCCCAACCCAAAAGTATTCTCTTTCTTAGTTTT	3529
QY	2161	CAGAAGTACTGCGATCACACGCGAGTTTACCTTTGGCGTGTGCCCTGTGTACCTGGCAG	2220
Db	3530	CAGAAGTACTGCGATCACACGCGAGTTTACCTTTGGCGTGTGCCCTGTGTACCTGGCAG	3589
QY	2221	AGAAGAGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG	2280
Db	3590	AGAAGAGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG	3649
QY	2281	CTATTTCCTTTAGACAGCGGACTGTATATAACAAGCCTTAACATTTGCTGCAAGATTTGCCCT	2340
Db	3650	CTATTTCCTTTAGACAGCGGACTGTATATAACAAGCCTTAACATTTGCTGCAAGATTTGCCCT	3709
QY	2341	CTTGAATT 2348	
Db	3710	CTTGAATT 3717	
RESULT	3		
ID	ABL39774		
XX	ABL39774 standard; cDNA; 5757 BP.		
AC	ABL39774;		
DT	10-MAY-2002 (first entry)		
DE	Human NS cDNA sequence SEQ ID NO:84.		
KW	Human; cytostatic; osteopathic; gynaecological; neuroprotective;		
KW	antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HI-		
KW	viral; vasotrophic; antiarteriosclerotic; antiinflammatory; dermatological;		
KW	anorectic; muscular; antifertility; cardiovascular; anticoagulant;		
KW	antifibrotic; hypotension; antiasthmatic; immunomodulator; cardiac;		
KW	anticongestant; antidiabetic; tranquilizer; antidepressant; neurolept		
KW	gastrointestinal; virucide; analgesic; cerebroprotective; nootropic;		
KW	contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;		
KW	rheumatoid arthritis; degenerative disease; multiple sclerosis; psoriasis;		
KW	endometriosis; catalect; restenosis; atherosclerosis; glaucoma;		
KW	inflammation; skin disorder; obesity; muscular dystrophy; AIDS;		
KW	infertility; cardiovascular disease; coagulation disease; hypertension		
KW	ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration		
KW	diabetes; anxiety; depression; schizophrenia; viral disease; stroke;		
OS	gastric ulcer; Alzheimer's disease; gene; ss.		
XX	Homo sapiens.		
XX	WO2000206315-A2.		
PX	24-JAN-2002.		
PD			

PF 17-JUL-2001; 2001WO-1L00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
XX (COMP-) COMPUGEN LTD.
PI Mintz L, Freilich S, Bernstein J;
XX
XX WPI: 2002-155037/20.
DR P-PSDB; ABB06120.
XX
XX One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
XX Claim 1; Page 124-126; 290pp; English.
XX
XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antisthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquiliser, antiulcer,
CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
XX Sequence 5757 BP; 1485 A; 1490 C; 1392 G; 1388 T; 2 other;

Query Match 92.9%; Score 2181.2; DB 24; Length 5757;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 2268; Conservative 1; Mismatches 4; Indels 75; Gaps 1;

QY 1 CCATGCGGGCCCTCACAGCCCGCGGGAGCCGAGCCCGCTGCCAGGCTGGCGCGG 60
DB 150 CCATGCGGGCCCTCCAGCCCGCGGGAGCCGCGCGCTGCCAGGCTGGCGCGG 209
QY 61 CSGTGCGCGATGTAGCGGGCTCCGAGATCCCGAGCTCTCCCTGCTCCCGTCTCTCGCGAT 120
DB 210 CCGTGGCGATGTAGCGGGCTCCGAGATCCCGAGCTCTCCCTGCTCCCGTCTCTCGCGAT 269
QY 121 CTCCTGACCGCTCTCCAGAGCCGAGCCCGGGGGTGGCCAGGCGCTGAGGGCCCT 180
DB 270 CTCCTGACCGCTCTCCAGAGCCGAGCCCGGGGGTGGCCAGGCGCTGAGGGCCCT 329
QY 181 GGGTCTGTATGCCCAAGCTCCCTCTCTGAGAGCCACAGCACCCAGACTTG 240
DB 330 GGGTCTGTATGCCCAAGCTCCCTCTCTGAGAGCCACAGCACCCAGACTTG 389
QY 241 GGGCAGCGCCAGGAGCGAGCTGGGCGAGTGGAGCCAGAGGGCCGAGCGCGGGG 300
DB 390 GGGCAGCGCCAGGAGCGAGCTGGGCGAGTGGAGCCAGAGGGCCGAGCGCGGGG 449
QY 301 CCACCAATGGCCCCAAGCCCTGCCCTGGCTCTGTGTGGATGGGCGCGGAGTGTGCGTG 360
DB 450 CCACCAATGGCCCCAAGCCCTGCCCTGGCTCTGTGTGGATGGGCGCGGAGTGTGCGTG 509
QY 361 CCACAGGCCACCCAGCAGCGATCCGGCTGCCCTGCCAGGCGCTGGGGGGGCCCCC 420
DB 510 CCACAGGCCACCCAGCAGCGATCCGGCTGCCCTGCCAGGCGCTGGGGGGGCCCCC 569
QY 421 TGGGGTGTGGCTGCCCGGGAGACCGAGAGCCCGAGAGCCCGCGCGGAGGGGCA 480

DB 570 TGGGGTGTGGCTGCCCGGGAGACCGAAGAGCCGAGGACCGCGCGGAGGGGCA 629
QY 481 GCCTTTGGAGATGGTGGACAACTGAGGGGCAAGTGGGGCAGGGCTACTACGTGAGA 540
DB 630 GCCTTTGGAGATGGTGGACAACTGAGGGGCAAGTGGGGCAGGGCTACTACGTGAGA 689
QY 541 TGACCGTGGGCGAGCCCGCGCAGACGCTCAACATCCTGGTGGATACAGGAGCAGTAAC 600
DB 690 TGACCGTGGGCGAGCCCGCGCAGACGCTCAACATCCTGGTGGATACAGGAGCAGTAAC 749
QY 601 TTGACGTGGTGTGTCGCCCGACCCCTTCCTGTGATCCTACTACAGGAGCAGTGTCCA 660
DB 750 TTGACGTGGTGTGTCGCCCGACCCCTTCCTGTGATCCTACTACAGGAGCAGTGTCCA 809
QY 661 GCACATACCGGGACCTCCGGAAGGTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAG 720
DB 810 GCACATACCGGGACCTCCGGAAGGTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAG 869
QY 721 GGGAGCTGGGCAACCGACCTGTAAGCATCCCGCATGGCCCCAACGTCACGTGCGTGCCA 780
DB 870 GGGAGCTGGGCAACCGACCTGTAAGCATCCCGCATGGCCCCAACGTCACGTGCGTGCCA 929
QY 781 ACATTCGTGCGCATCACTGAATCAGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGCA 840
DB 930 ACATTCGTGCGCATCACTGAATCAGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGCA 989
QY 841 TCCTGGGGCTGGCTATGCTGAGATTGCCAGGCGCTGAGGACTCCCTGGAGCCTTTCTTG 900
DB 990 TCCTGGGGCTGGCTATGCTGAGATTGCCAG----- 1020
QY 901 ACTCTCTGTTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGCTGTG 960
DB 1021 -----GCTTTGTGCTGTG 1034
QY 961 GCTTCCCGCTCAACAGTCTGAACTGTGGCTCTGTGGAGGAGGATGATCATTTGGAG 1020
DB 1035 GCTTCCCGCTCAACAGTCTGAACTGTGGCTCTGTGGAGGAGGATGATCATTTGGAG 1094
QY 1021 GTATCGACCACTCGCTGTACAGGCGAGTCTCTGCTATACACCCATCCGGCGGAGTGGT 1080
DB 1095 GTATCGACCACTCGCTGTACAGGCGAGTCTCTGCTATACACCCATCCGGCGGAGTGGT 1154
QY 1081 ATTATGAGGTGATCATTTGTGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCA 1140
DB 1155 ATTATGAGGTGATCATTTGTGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCA 1214
QY 1141 AGGAGTACAATATGACAAGAGCATTTGGACACTGGCACACCAACCTTCCTTTGGCCCA 1200
DB 1215 AGGAGTACAATATGACAAGAGCATTTGGACACTGGCACACCAACCTTCCTTTGGCCCA 1274
QY 1201 AGAAGTGTTTGAAGCTGCAGTCAATCCATCAAGGCGAGCTCTCCAGGAGAGTTC 1260
DB 1275 AGAAGTGTTTGAAGCTGCAGTCAATCCATCAAGGCGAGCTCTCCAGGAGAGTTC 1334
QY 1261 CTGATGCTTTCTGGCTAGGAGCAGCTGGTGTGCTGGCAAGCAGGACACCCCTTTGGA 1320
DB 1335 CTGATGCTTTCTGGCTAGGAGCAGCTGGTGTGCTGGCAAGCAGGACACCCCTTTGGA 1394
QY 1321 ACATTTTCCCAGTCACTCTACCTTAATGGGTGAGGTTTACCAACAGCTCTTCCGCA 1380
DB 1395 ACATTTTCCCAGTCACTCTACCTTAATGGGTGAGGTTTACCAACAGCTCTTCCGCA 1454
QY 1381 TCACCATCTTTCCGAGCAATACCTGGGCGCAGTGGAGATGTGGCCACGCTCCCAAGACG 1440
DB 1455 TCACCATCTTTCCGAGCAATACCTGGGCGCAGTGGAGATGTGGCCACGCTCCCAAGACG 1514
QY 1441 ACTGTTACAAGTTTGGCCATCTCAGTCATCCACGGCAGCTGTTATGGAGCTGTATCA 1500
DB 1515 ACTGTTACAAGTTTGGCCATCTCAGTCATCCAGGCGACCTGTTATGGAGCTGTATCA 1574
QY 1501 TGGAGGCTCTTACGTTGTCTTTGATCGGGCCGGAACAAATTTGGCTTGTCTGTCAAGCG 1560


```
Db 421 CTGGCCACCGACCTGGTAAAGCATCCCATGGCCCAACGCTCACTGTGCGTGCCAAACATT 480
QY 786 GCTCCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCCAACTGGGAAGGCATCCG 845
Db 481 GCTCCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCCAACTGGGAAGGCATCCG 540
QY 846 GGGTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 905
Db 541 GGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 600
QY 906 CTGCTAAAGCAGACCCAGCTTCCCAACCTCTCTCTCCCTGCGAGCTTTGTGTGCTGGCTTC 965
Db 601 CTGCTAAAGCAGACCCAGCTTCCCAACCTCTCTCTCCCTGCGAGCTTTGTGTGCTGGCTTC 660
QY 966 CCCCTCAACCAAGTCTGAAAGTGGCTGCTCTGCGAGGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCCTCAACCAAGTCTGAAAGTGGCTGCTCTGCGAGGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACCTCGCTGTACACAGGCACTCTCTGATATACACCCATCCGGGGAGTGGTATTAT 1085
Db 721 GACCACCTCGCTGTACACAGGCACTCTCTGATATACACCCATCCGGGGAGTGGTATTAT 780
QY 1086 GAGGTGATCATTTGCGGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 1145
Db 781 GAGGTGATCATTTGCGGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 840
QY 1146 TACAACATATGACAAGAGCATTTGTGACAGTGGCCACCAACCACTTCGTTTGGCCCAAGAAA 1205
Db 841 TACAACATATGACAAGAGCATTTGTGACAGTGGCCACCAACCACTTCGTTTGGCCCAAGAAA 900
QY 1206 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCAAGCTCTCCACGGAGAAAGTTCCCTGAT 1265
Db 901 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCAAGCTCTCCACGGAGAAAGTTCCCTGAT 960
QY 1266 GGTTCCTGGCTAGGAGACAGCTGTGTGCTGCGAAGCAGGACACCCCTTTGGAACATT 1325
Db 961 GGTTCCTGGCTAGGAGACAGCTGTGTGCTGCGAAGCAGGACACCCCTTTGGAACATT 1020
QY 1326 TTCCCACTGATCTCACTCTACCTAATAGGTGAGTTTACCAACCACTCTTCCCGCATCAC 1385
Db 1021 TTCCCACTGATCTCACTCTACCTAATAGGTGAGTTTACCAACCACTCTTCCCGCATCAC 1080
QY 1386 ATCCTTCGCGACGAATACCTCGGCGCAGTGGAAAGATGTGGCCACGTCCTCCCAAGACGACTGT 1445
Db 1081 ATCCTTCGCGACGAATACCTCGGCGCAGTGGAAAGATGTGGCCACGTCCTCCCAAGACGACTGT 1140
QY 1446 TACAAGTTTGGCATCTCAGATCATCCAGCGGCACTGTTATGGGAGCTGTATCATCGAG 1505
Db 1141 TACAAGTTTGGCATCTCAGATCATCCAGCGGCACTGTTATGGGAGCTGTATCATCGAG 1200
QY 1506 GGCCTTACGTTGCTTTGATCGGGCCCCGAAACGAATTTGGCTTTGCTGCAGCGCTTGC 1565
Db 1201 GGCCTTACGTTGCTTTGATCGGGCCCCGAAACGAATTTGGCTTTGCTGCAGCGCTTGC 1260
QY 1566 CATGTGCACGATAGTTTCAGGACGGCAGCGGTGGAAAGGCCCTTTTGTACCTTGGACATG 1625
Db 1261 CATGTGCACGATAGTTTCAGGACGGCAGCGGTGGAAAGGCCCTTTTGTACCTTGGACATG 1320
QY 1626 GAAGACTGTGGCTACAACATTCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1685
Db 1321 GAAGACTGTGGCTACAACATTCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1380
QY 1686 GTCATGGCTGCGATCTGGCCCTCTTTCATGCTGCCACTCTGCCTCATGTTGTGTGTCAGTGG 1745
Db 1381 GTCATGGCTGCGATCTGGCCCTCTTTCATGCTGCCACTCTGCCTCATGTTGTGTGTCAGTGG 1440
QY 1746 CGTGGCTCCGCTGCCTCGCGCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1805
Db 1441 CGTGGCTCCGCTGCCTCGCGCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGGAGGCCCCATGGGCAAGCATGAGATTCCCTTGGACCAACCTCCCGTGGTTCA 1865
```

```
Db 1501 AAGTGGAGGCCCCATGGGCAAGCATGAGATTCCCTTGGACCAACCTCCCGTGGTTCA 1560
QY 1866 CTTTGGTTCACAAGTAGGAGACACACATGGCACCTGTGGCCAGAGCACTCTCAGAACCTCC 1925
Db 1561 CTTTGGTTCACAAGTAGGAGACACACATGGCACCTGTGGCCAGAGCACTCTCAGAACCTCC 1620
QY 1926 CCACCCCAACCAATGCCCTTCTGGCTTTGATGGAGAGGAAAAGGCTGGCAAGCTGGTTCAG 1985
Db 1621 CCACCCCAACCAATGCCCTTCTGGCTTTGATGGAGAGGAAAAGGCTGGCAAGCTGGTTCAG 1680
QY 1986 GGAATGTACCTGTAGGAAACAGAAAGAGAGAAAGCACTCTGCTGGCGGAATACT 2045
Db 1681 GGAATGTACCTGTAGGAAACAGAAAGAGAGAAAGCACTCTGCTGGCGGAATACT 1740
QY 2046 CTTTGGTCAACCTCAAAATTTAAGTCGCGGAAATTTCTGCTTTGAAACTTCAGCCCTGAACCT 2105
Db 1741 CTTTGGTCAACCTCAAAATTTAAGTCGCGGAAATTTCTGCTTTGAAACTTCAGCCCTGAACCT 1800
QY 2106 TTGTCCCACTATCCTTTAAATTTCTCAACCCCAAGTAATTTCTTTCTTTAGTTTCAGAA 2165
Db 1801 TTGTCCCACTATCCTTTAAATTTCTCAACCCCAAGTAATTTCTTTCTTTAGTTTCAGAA 1860
QY 2166 GTACTGGCATCACAGCAGGTTTACCTTGGCGTGTGTCCCTGTGCTACCCCTGGCAGAGAG 2225
Db 1861 GTACTGGCATCACAGCAGGTTTACCTTGGCGTGTGTCCCTGTGCTACCCCTGGCAGAGAG 1920
QY 2226 AGACCAAGCTTGTTCCTCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGTCTATT 2285
Db 1921 AGACCAAGCTTGTTCCTCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGTCTATT 1980
QY 2286 TGCTTTAGAGACAGGAGCTGTATAAACAAGCCTTACATTTGGTGCAAGATTGGCTCTTTGA 2345
Db 1981 TGCTTTAGAGACAGGAGCTGTATAAACAAGCCTTACATTTGGTGCAAGATTGGCTCTTTGA 2040
QY 2346 ATT 2348
Db 2041 ATT 2043

RESULT 7
AASL1517
ID AASL1517 standard; cDNA: 2070 BP.
XX
AC AASL1517;
XX
DT XX
XX 24-OCT-2001 (first entry)
XX Human cDNA encoding Aspartyl protease 2(a), Asp2(a).
DE Human: Aspartyl protease; Asp2(a): beta-secretase; nototropic;
KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
XX amyloid-beta; Abeta; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..1506
FT /*tag= a
FT /product= "Asp2(a)"
FT sig_peptide 1..63
FT /*tag= b
FT sig_peptide 64..135
FT /*tag= c
FT sig_peptide 136..171
FT /*tag= d
FT mat_peptide 172..1503
FT /*tag= e
FT /label= Mature_Asp2(a)
XX
PN WO200149098-A2.
XX
```


QY 1686 GTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG 1745
Db 1381 GTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG 1440
QY 1746 CGCTGCCCTCCGCTGCTGGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGCCCTCCGCTGCTGGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGAGGAGGCCCTGGGAGAGATAGAGATTCCCTGGACCACACCTCCGCTGGTTCA 1865
Db 1501 AAGTGAGGAGGCCCTGGGAGAGATAGAGATTCCCTGGACCACACCTCCGCTGGTTCA 1560
QY 1866 CTTTGGTACAGTAGGAGACACAGATGGACCTGTGGCCAGACACCTCAGGACCCCTCC 1925
Db 1561 CTTTGGTACAGTAGGAGACACAGATGGACCTGTGGCCAGACACCTCAGGACCCCTCC 1620
QY 1926 CCACCCACCAATGCTCTGCTGCTTGATGAGAGAGAAAGGCTGGCAAGCTGGGTTCCAG 1985
Db 1621 CCACCCACCAATGCTCTGCTGCTTGATGAGAGAGAAAGGCTGGCAAGCTGGGTTCCAG 1680
QY 1986 GGACTGTACCTGTAGGAACAGAAAGAGAAAGAAAGCACTCTGCTGGCGGGAATACT 2045
Db 1681 GGACTGTACCTGTAGGAACAGAAAGAGAAAGAAAGCACTCTGCTGGCGGGAATACT 1740
QY 2046 CTTGGTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAAGCCCTGAACT 2105
Db 1741 CTTGGTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAAGCCCTGAACT 1800
QY 2106 TTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGATTTCTTCTTTAGTTTCAGAA 2165
Db 1801 TTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGATTTCTTCTTTAGTTTCAGAA 1860
QY 2166 GTACTGGCATCACAGCGAGTTACCTTGGCGTGTGCTCCCTGTGTACCTGGCAGAGAG 2225
Db 1861 GTACTGGCATCACAGCGAGTTACCTTGGCGTGTGCTCCCTGTGTACCTGGCAGAGAG 1920
QY 2226 AGACCAAGCTGTTTCCCTGCTGGCCAAAGTAGTAGGAGAGGATGCACAGTTTGTCTATT 2285
Db 1921 AGACCAAGCTGTTTCCCTGCTGGCCAAAGTAGTAGGAGAGGATGCACAGTTTGTCTATT 1980
QY 2286 TGCPTTAGAGACAGGCACTGTATAACAAGCCTTAACATGCTGCAAGATTGCCCTTTGA 2345
Db 1981 TGCPTTAGAGACAGGCACTGTATAACAAGCCTTAACATGCTGCAAGATTGCCCTTTGA 2040
QY 2346 ATT 2348
Db 2041 ATT 2043

RESULT 8

AAS11702
ID AAS11702 standard; DNA: 2070 BP.

XX AAS11702;

AC AAS11702;

XX 24-Oct-2001 (first entry)

DE DNA encoding human aspartyl protease 2a (Asp-2a).

XX Human: aspartyl protease 1; Asp-1; nootropic; neuroprotective;
KW aspartyl protease 2; Asp2; amyloid protein precursor; APP;
KW beta-secretase; Alzheimer's disease; ds.
XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 1..1506
FT /*tag= a
FT /product= "Aspartyl protease-2a (Asp-2a)"
FT 1..63
FT sig_peptide /*tag= b
FT 64..135
FT misc_feature /*tag= c

FT misc_feature /note= "Pre-propeptide"
FT 136..171
FT /*tag= d
FT /note= "Propeptide"
FT 172..1503
FT /*tag= e
FT /note= "Mature Aspartyl protease-2a"

XX WO200149097-A2.

XX 12-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00797.

XX 09-MAY-2001; 2001WO-IB00797.

XX (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.

XX

PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

XX WPI; 2001-502548/55.

DR P-PSDB; AAU07102.

XX

PT Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta

PT secretase activity of Asp2 useful for identifying inhibitors of Asp2

PT activity -

XX

PS Claim 98; Fig 2; 185pp; English.

XX

CC The invention relates to a novel purified polypeptide comprising a

CC fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the

CC Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide

CC and the fragment retain the beta-secretase activity of the mammalian Asp2

CC protein. Also included is an isoform of amyloid protein precursor (APP)

CC comprising the amino acid sequence of a APP or its fragment containing

CC an APP cleavage site recognisable by a mammalian beta-secretase, and

CC further comprising two lysine residues at the carboxyl terminus of the

CC amino acid sequence of the mammalian APP or APP fragment. The

CC polypeptides are used for assaying for modulators of beta-secretase

CC activity; identifying agents that inhibit the APP processing activity

CC of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that

CC modulate the activity of Asp2; and for reducing cellular production of

CC amyloid beta (Abeta) from APP. Agents identified by the above methods

CC are useful for treating Alzheimer's disease; and for identifying

CC modulators of amyloid-beta (Abeta) peptide production, for use in

CC designing therapeutics for the treatment or prevention of Alzheimer's

CC disease. Probes and primers derived from Asp nucleic acid sequences

CC are useful for detecting Hu-Asp nucleic acids in in vitro assays and in

CC Northern and Southern blots. The present sequence represents the

CC coding sequence of human Asp-2a used in the methods of the invention.

XX

SO Sequence 2070 BP; 476 A; 583 C; 562 G; 449 T; 0 other;

Query Match 86.9%; Score 2039.8; DB 22; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 306 ATGGCCCAAGCCCTGCCCTGGCTCCTCTGTGGATGGCGGGAGTGTGCTGCTGCCAC 365
Db 1 ATGGCCCAAGCCCTGCCCTGGCTCCTCTGTGGATGGCGGGAGTGTGCTGCTGCCAC 60

QY 366 GGCACCCACGACGGCATCCGGCTGCCCTGTGCGACGGCCCTGGGGGGGCCCCCTCTGGGG 425
Db 61 GGCACCCACGACGGCATCCGGCTGCCCTGTGCGACGGCCCTGGGGGGGCCCCCTCTGGGG 120
QY 426 CTGGCGCTGCCCGGGGAGACCCGAGAGAGCCCGGAGGCGCGGAGGCGACGCTTT 485
Db 121 CTGGCGCTGCCCGGGGAGACCCGAGAGAGCCCGGAGGCGCGGAGGCGGAGGCGACGCTTT 180


```
QY 506 GTGGTGTGTCGCCCCACCCCTTCTGTCATCGCTACACAGAGGAGCGAGTGTCCACACA 665
DB 301 CTGGTGTGTCGCCCCACCCCTTCTGTCATCGCTACACAGAGGAGCGAGTGTCCACACA 360
QY 666 TACCGGAGACCTCCGGAAGGTTGTATGTGTCCTTACACCCAGGCAAGTGGGAAGGGAG 725
DB 361 TACCGGAGACCTCCGGAAGGTTGTATGTGTCCTTACACCCAGGCAAGTGGGAAGGGAG 420
QY 726 CTGGGACCCGACCTGTAAAGCATCCCCCATGGGCCCAACGTCACCTGTGCGTGCCCAACATT 785
DB 421 CTGGGACCCGACCTGTAAAGCATCCCCCATGGGCCCAACGTCACCTGTGCGTGCCCAACATT 480
QY 786 GCTGCCATCTAGTATCAGACAAAGTTCTTCAACAACGGCTCCAACTGGGAAGGCATCCG 845
DB 481 GCTGCCATCTAGTATCAGACAAAGTTCTTCAACAACGGCTCCAACTGGGAAGGCATCCG 540
QY 846 GGGCTGGGCTATGCTCAGATTGCGACGGCTGACGACTCCCTGGAGCCCTTCTTTCACCTCT 905
DB 541 GGGCTGGGCTATGCTCAGATTGCGACGGCTGACGACTCCCTGGAGCCCTTCTTTCACCTCT 600
QY 906 CTGTTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAAGCTTGTGCTGCTGCTTC 965
DB 601 CTGTTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAAGCTTGTGCTGCTGCTTC 660
QY 966 CCCCTCAACCACTGTGAAGTGTGGGCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC 1025
DB 661 CCCCTCAACCACTGTGAAGTGTGGGCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACCTCGCTACACAGGCGAGTCTGTGTATACACCCATCCCGGGAGTGGTATTAT 1085
DB 721 GACCACCTCGCTACACAGGCGAGTCTGTGTATACACCCATCCCGGGAGTGGTATTAT 780
QY 1086 GAGGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 1145
DB 781 GAGGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 840
QY 1146 TACAACATGACAAAGACATTTGGACAGTGGGACACCAACCACTTCGTTTGGCCCAAGAA 1205
DB 841 TACAACATGACAAAGACATTTGGACAGTGGGACACCAACCACTTCGTTTGGCCCAAGAA 900
QY 1206 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCGAGCCCTCTCCACGGGAGAAATTCCTGTAT 1265
DB 901 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGCGAGCCCTCTCCACGGGAGAAATTCCTGTAT 960
QY 1266 GGTTCCTGGCTAGGAGAGCAGTGTGTGTGCGAAGCAGCAGCACCCCTTGGAAACATT 1325
DB 961 GGTTCCTGGCTAGGAGAGCAGTGTGTGTGCGAAGCAGCAGCACCCCTTGGAAACATT 1020
QY 1326 TTCCCAAGTCATCTACTCTACCTAATGGGTGAGGTTACCAACCACTCCTTCGGCATCACC 1385
DB 1021 TTCCCAAGTCATCTACTCTACCTAATGGGTGAGGTTACCAACCACTCCTTCGGCATCACC 1080
QY 1386 ATCCTTCGGCAGCAATACCTGCGGCAGTGGGAAGATGTGGCCACGTGCCAAGACGACTGT 1445
DB 1081 ATCCTTCGGCAGCAATACCTGCGGCAGTGGGAAGATGTGGCCACGTGCCAAGACGACTGT 1140
QY 1446 TACAAGTTTGCCATCTCAGTCAATCCACGGGCACTGTTATGGAGCTGTTATCATGAGG 1505
DB 1141 TACAAGTTTGCCATCTCAGTCAATCCACGGGCACTGTTATGGAGCTGTTATCATGAGG 1200
QY 1506 GGCTTCACGTTGCTTTGATCGGGCCGAAAACAGAAATGGCTTTGCTGTGTCAGGCTTGC 1565
DB 1201 GGCTTCACGTTGCTTTGATCGGGCCGAAAACAGAAATGGCTTTGCTGTGTCAGGCTTGC 1260
QY 1566 CATGTGACGATGAGTTTCAGGAGCGCAGCGTGGGAAGGCCCTTTTGTCAACCTTGGACATG 1625
DB 1261 CATGTGACGATGAGTTTCAGGAGCGCAGCGTGGGAAGGCCCTTTTGTCAACCTTGGACATG 1320
QY 1626 GAAGACTGTGGCTACAACATTCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
DB 1321 GAAGACTGTGGCTACAACATTCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380
QY 1686 GTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTCTGCCCTCATGCTGTGTCAGTGG 1745

Db 1381 GTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGG 1440
QY 1746 CGCTGCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTG 1805
DB 1441 CGCTGCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTAGGAGGAGCCATGGGAGAGATAGAGATCCCTGGAGCAACACACCTCCGTTGGTTCA 1865
DB 1501 AAGTAGGAGGAGCCATGGGAGAGATAGAGATCCCTGGAGCAACACCTCCGTTGGTTCA 1560
QY 1866 CTTTGTCTACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCACTCTCAGGACCCCTCC 1925
DB 1561 CTTTGTCTACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCACTCTCAGGACCCCTCC 1620
QY 1926 CCACCCACCAATGCGCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGGTGGGTTCCAG 1985
DB 1621 CCACCCACCAATGCGCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGGTGGGTTCCAG 1680
QY 1986 GGACTGTACTCTAGGAAACAGAAAGAGAAAGAAAGCACTCTCTGCTGGCGGGAATACT 2045
DB 1681 GGACTGTACTCTAGGAAACAGAAAGAGAAAGAAAGCACTCTCTGCTGGCGGGAATACT 1740
QY 2046 CTTGTTCACTCTCAAAATTTAAGTCGGGAAATTTCTGCTGTGTGAAACTTCAAGCCCTGAAACCT 2105
DB 1741 CTTGTTCACTCTCAAAATTTAAGTCGGGAAATTTCTGCTGTGTGAAACTTCAAGCCCTGAAACCT 1800
QY 2106 TTGTTCCACCATCTCTTAAATTTCCACCCCAAAAGTATTTCTTTCTTAGTTTCAGAA 2165
DB 1801 TTGTTCCACCATCTCTTAAATTTCCACCCCAAAAGTATTTCTTTCTTAGTTTCAGAA 1860
QY 2166 GTACTGGCATCAACGAGGTTTACCTTGGCGTGTCTCCCTGTGGTACCCTGGCAGACAAG 2225
DB 1861 GTACTGGCATCAACGAGGTTTACCTTGGCGTGTCTCCCTGTGGTACCCTGGCAGACAAG 1920
QY 2226 AGACCAAGCTTGTTCCTGCTGCGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTATT 2285
DB 1921 AGACCAAGCTTGTTCCTGCTGCGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTATT 1980
QY 2286 TGCCTTTAGAGACAGGACTGTATAACAAGCTTAACATTGGTGCAAGATTGCTCTTGA 2345
DB 1981 TGCCTTTAGAGACAGGACTGTATAACAAGCTTAACATTGGTGCAAGATTGCTCTTGA 2040
QY 2346 ATT 2348
DB 2041 ATT 2043

RESULT 11
ABL49914
ID ABL49914 standard; DNA; 2070 BP.
XX
AC ABL49914;
XX
DT 31-MAY-2002 (first entry)
XX
DE Human aspartyl protease nucleotide sequence SEQ ID NO:1.
XX
KW Beta-secretase; enzyme; cleavage site; amyloid protein precursor; APP;
KW aspartyl protease; neuroprotective; nootropic; beta-secretase inhibitor;
XX Alzheimer's disease; gene; ds.
OS Homo sapiens.
XX
PN WO200206306-A2.
XX
PD 24-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-US23035.
XX
PR 19-JUL-2000; 2000US-219795P.
XX 12-MAR-2001; 2001US-275251P.
XX
```

PA	(PHAA) PHARMACIA & UPJOHN CO.
XX	
PI	Yan R, Tomasselli AG, Gurney ME, Emmons TL, Bienkowski MJ;
PI	Heinrikson RL;
XX	
XX	
DR	WPI: 2002-216995/27.
DR	P-PSDB: ABB06409.
XX	
XX	
PT	Novel substrates for human aspartyl protease useful for identifying
PT	modulators of beta secretase activity of aspartyl protease for treating
PT	Alzheimer's disease -
XX	
XX	Claim 1; Page 117; 188pp; English.
XX	
CC	The present invention describes an isolated peptide (I) comprising a
CC	sequence of at least four amino acids, where the peptide is a substrate
CC	for conducting aspartyl protease assays. (I) has neuroprotective and
CC	neotropic activities, and can be used as an inhibitor of beta-secretase
CC	activity. A beta-secretase modulator from the present invention can be
CC	used for inhibiting beta-secretase activity in vivo, and in the
CC	manufacture of a medicament for the treatment of Alzheimer's disease.
CC	Pharmaceutical compositions from the present invention can be used for
CC	treating a disease or condition characterised by an abnormal beta-
CC	secretase activity. (I) is useful for identifying agents that modulate
CC	the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful
CC	as a core structure to construct derivatives. ABL49914 to ABL49925 and
CC	ABB06409 to ABB06593 represent sequences used in the exemplification
CC	of the present invention.


```
XX 04-JAN-2001.
XX 27-JUN-2000; 2000WO-US17661.
XX
XX 28-JUN-1999; 99US-0141363.
PR 25-JAN-1999; 99US-0168060.
PR 25-JAN-2000; 2000US-0177836.
PR 27-JAN-2000; 2000US-0178368.
PR 08-JUN-2000; 2000US-0210292.
XX
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
XX Tang JJN, Lin X, Koelsch G;
XX WPI; 2001-102885/11.
XX
XX Purified recombinant catalytically active memapsin 2, used to screen
XX inhibitors of it, which are used to treat and prevent Alzheimer's
XX disease -
XX
XX Example 1; Page 71-72; 86pp; English.
XX
XX The present invention relates to a purified recombinant
XX catalytically active memapsin 2. The invention may be used for
XX isolating inhibitors which are used to treat or prevent
XX Alzheimer's disease. The invention may also be used to screen
XX for individuals more genetically prone to develop Alzheimer's
XX disease.
XX
XX Sequence 3252 BP; 804 A; 863 C; 811 G; 771 T; 3 other;
XX
XX Query Match 85.3%; Score 2004; DB 22; Length 3252;
XX Best Local Similarity 100.0%; Pred No. 0;
XX Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 345 GCGGGAGTGTGCTGCCACGACACCCAGCAGCGCATCCGGCTGCCCTGCGCAGCGGC 404
XX 1 GCGGGAGTGTGCTGCCACGACACCCAGCAGCGCATCCGGCTGCCCTGCGCAGCGGC 60
XX
XX 405 CTGGGGGGCCCCCTGGGGCTGCGGCTGCCCGGGAGACCGACGAGCCCGGAGGAG 464
XX
XX 61 CTGGGGGGCCCCCTGGGGCTGCGGCTGCCCGGGAGACCGACGAGCCCGGAGGAG 120
XX
XX 465 CCGCGCGGAGGGGACGCTTTGTGGAGATGCTGGACAACTGAGGGGCAAGTGGGGCAG 524
XX 121 CCGCGCGGAGGGGACGCTTTGTGGAGATGCTGGACAACTGAGGGGCAAGTGGGGCAG 180
XX
XX 525 GGCTACTACGTGGAGATGACCGTGGGACGCCCCCGGACGCGCTCAACATCCTGGTGGAT 584
XX 181 GGCTACTACGTGGAGATGACCGTGGGACGCCCCCGGACGCGCTCAACATCCTGGTGGAT 240
XX
XX 585 ACAGGACGAGTAACCTTGGAGTGGGTGCTGCCCGGACGCCCCCGGACGCGCTCAACATCCTGGTGGAT 644
XX 241 ACAGGACGAGTAACCTTGGAGTGGGTGCTGCCCGGACGCCCCCGGACGCGCTCAACATCCTGGTGGAT 300
XX
XX 645 CAGAGCAGCTGTCAGCACATACCGGACCTCCGGAGGAGTGTATGTCGCCCTACACC 704
XX 301 CAGAGCAGCTGTCAGCACATACCGGACCTCCGGAGGAGTGTATGTCGCCCTACACC 360
XX
XX 705 CAGGGCAAGTGGGAAGGGAGCTGGGACCGACCTGTAAGCATGCCCGCATGGCCCCAAC 764
XX 361 CAGGGCAAGTGGGAAGGGAGCTGGGACCGACCTGTAAGCATGCCCGCATGGCCCCAAC 420
XX
XX 765 GTCAGTGTGGTGCACATTTGTGGCATCACTGAATCAGACAAGTTCCTATCAACGGC 824
XX 421 GTCAGTGTGGTGCACATTTGTGGCATCACTGAATCAGACAAGTTCCTATCAACGGC 480
XX
XX 825 TCCAACCTGGGAAGGCATCCTGGGGCTGGCCTATGCTAGATGCCAGGCTGAGACTCC 884
XX 481 TCCAACCTGGGAAGGCATCCTGGGGCTGGCCTATGCTAGATGCCAGGCTGAGACTCC 540
XX
XX 885 CTGGAGCCTTTCTTGACTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTG 944
```

```
Db 541 CTGGAGCCTTTCTTGACTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTG 600
Qy 945 CAGCTTTGTGGTGTGCTGGCTTCCCCCTCAACGAGTCTCAAGTGTGGCTCTGTGGAGGG 1004
Db 601 CAGCTTTGTGGTGTGCTGGCTTCCCCCTCAACGAGTCTCAAGTGTGGCTCTGTGGAGGG 660
Qy 1005 AGCATGATCAATTTGGAGGTATCGACCACTCGCTGTACACAGCAGCTCTCTGGTATACACC 1064
Db 661 AGCATGATCAATTTGGAGGTATCGACCACTCGCTGTACACAGCAGCTCTCTGGTATACACC 720
Qy 1065 ATCCGGGGGAGTGGTATTATGAGGTGATCATTTGTGGGGTGGAGATCAATTTGGACAGGAT 1124
Db 721 ATCCGGGGGAGTGGTATTATGAGGTGATCATTTGTGGGGTGGAGATCAATTTGGACAGGAT 780
Qy 1125 CTGAAATGGACTGCAAGGAGTACAACATATGACAAGAGCATTTGGGACAGTGGCACACC 1184
Db 781 CTGAAATGGACTGCAAGGAGTACAACATATGACAAGAGCATTTGGGACAGTGGCACACC 840
Qy 1185 AACCTTCGTTTGGCCAAAGAAAGTGTTTGAAGTGCAGTCAAAATCCATCAAGGACGCTCC 1244
Db 841 AACCTTCGTTTGGCCAAAGAAAGTGTTTGAAGTGCAGTCAAAATCCATCAAGGACGCTCC 900
Qy 1245 TCCACGGAGAAAGTTCCCTGATGGTCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 1304
Db 901 TCCACGGAGAAAGTTCCCTGATGGTCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 960
Qy 1305 GGACACACCCCTTGGAAACATTTCCAGTGCATCTCACTCTACCTAATGGGTGAGGTACC 1364
Db 961 GGACACACCCCTTGGAAACATTTCCAGTGCATCTCACTCTACCTAATGGGTGAGGTACC 1020
Qy 1365 AACAGTCCCTTCGCGATCAGCATCTTCGCGAGCAATACCTGCGGCGAGTGAAGATGTG 1424
Db 1021 AACAGTCCCTTCGCGATCAGCATCTTCGCGAGCAATACCTGCGGCGAGTGAAGATGTG 1080
Qy 1425 GCCACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTGCATCCACGGGCACTGTT 1484
Db 1081 GCCACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTGCATCCACGGGCACTGTT 1140
Qy 1485 ATGGGAGCTGTTATCATGGAGGGCTTCTAGCTTCTTGTATCGGCGCCGAAACGAATTT 1544
Db 1141 ATGGGAGCTGTTATCATGGAGGGCTTCTAGCTTCTTGTATCGGCGCCGAAACGAATTT 1200
Qy 1545 GGGTTTCTGTACGCGCTTGCCATGTGCAGATGATTCAGGACGCGAGCGGTGGAAGGC 1604
Db 1201 GGGTTTCTGTACGCGCTTGCCATGTGCAGATGATTCAGGACGCGAGCGGTGGAAGGC 1260
Qy 1605 CCTTTTCTACCTTTGGACATGGAAGACTGTGGCTTACAACATTCACACAGACAGATGAGTCA 1664
Db 1261 CCTTTTCTACCTTTGGACATGGAAGACTGTGGCTTACAACATTCACACAGACAGATGAGTCA 1320
Qy 1665 ACCCTCATGACCATAGCCTATGTGCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTC 1724
Db 1321 ACCCTCATGACCATAGCCTATGTGCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTC 1380
Qy 1725 TGCCCTATGTGTGTCAGTGGCGCTGCTCCGCTGCTGCGCCAGCAGCAGCATGATGACTTT 1784
Db 1381 TGCCCTATGTGTGTCAGTGGCGCTGCTCCGCTGCTGCGCCAGCAGCAGCATGATGACTTT 1440
Qy 1785 GCTGATGACATCTCCCTGCTGAGTGGAGGGCCCATGGGCGAGAGATAGAGATTCCCTCT 1844
Db 1441 GCTGATGACATCTCCCTGCTGAGTGGAGGGCCCATGGGCGAGAGATAGAGATTCCCTCT 1500
Qy 1845 GGACACACCTCCCTGGTTCATTTGCTGTCACAGTAGAGACAGATGGCACCTGTGGC 1904
Db 1501 GGACACACCTCCCTGGTTCATTTGCTGTCACAGTAGAGACAGATGGCACCTGTGGC 1560
Qy 1905 CAGAGCAGCTCAGGACCCCTCCCGACCCACCAATGCTCTGCTGCTGATGGAGAGAGAAA 1964
Db 1561 CAGAGCAGCTCAGGACCCCTCCCGACCCACCAATGCTCTGCTGCTGATGGAGAGAGAAA 1620
Qy 1965 GGTGGCAAGTGGGTTCAGGGGACTGTACCTGTAGGAAACAGAAAAGAGAAAGAG 2024
```


Qy 1125 CTGAAATGGACTGCAAGGAGTCAAACTATGACAAGAGCATTTGGACAGTGGCACCACC 1184
Db 781 CTGAAATGGAGTGCAGAGGATGACAACTATGACAAGAGCATTTGGACAGTGGCACCACC 840
Qy 1185 AACCTTCCTTTTGGCCCAAGAAAGTGTGGAAGCTGCAGTCAAACTCCATCCATCAAGGCACCTCC 1244
Db 841 AACCTTCCTTTTGGCCCAAGAAAGTGTGGAAGCTGCAGTCAAACTCCATCCATCAAGGCACCTCC 900
Qy 1245 TCCACGGAGAAGTTCCTGTATGGTTCCTGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 1304
Db 901 TCCACGGAGAAGTTCCTGTATGGTTCCTGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 960
Qy 1305 GGCACACCCCTTGGAAACATTTTCCAGTCATCTACCTACCTAATGGGTGAGSTTACC 1364
Db 961 GGCACACCCCTTGGAAACATTTTCCAGTCATCTACCTACCTAATGGGTGAGSTTACC 1020
Qy 1365 AACCAGTCTTTCGGCATCACCATCTTCCGAGCAATACCTGCGGCCAGTGGGAAGATGTG 1424
Db 1021 AACCAGTCTTTCGGCATCACCATCTTCCGAGCAATACCTGCGGCCAGTGGGAAGATGTG 1080
Qy 1425 GCCAGTCCCAAGAGCAGCTGTACAAAGTTTGCCATCTCACAGTCAATCCACGGGCACCTGTT 1484
Db 1081 GCCAGTCCCAAGAGCAGCTGTACAAAGTTTGCCATCTCACAGTCAATCCACGGGCACCTGTT 1140
Qy 1485 ATGGAGCTGTTATCATGGAGGGCTTCTACGTTGCTTTGATCGGGCCCGAAACGAATT 1544
Db 1141 ATGGAGCTGTTATCATGGAGGGCTTCTACGTTGCTTTGATCGGGCCCGAAACGAATT 1200
Qy 1545 GCGTTTGTGTACGGCTTGCATGTGCACATGAGTTTCAGGACGGCAGCGGTGGAAAGC 1604
Db 1201 GCGTTTGTGTACGGCTTGCATGTGCACATGAGTTTCAGGACGGCAGCGGTGGAAAGC 1260
Qy 1605 CTTTGTGTCACCTTGGACATGGAAGACTGTGCTACAACTTCCACAGACAGATGAGTCA 1664
Db 1261 CTTTGTGTCACCTTGGACATGGAAGACTGTGCTACAACTTCCACAGACAGATGAGTCA 1320
Qy 1665 ACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTC 1724
Db 1321 ACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTC 1380
Qy 1725 TGCCTCATGGTGTGTCAGTGGCGCTGCGCTGCGCTGCGCCAGCAGCATGATGACTTT 1784
Db 1381 TGCCTCATGGTGTGTCAGTGGCGCTGCGCTGCGCTGCGCCAGCAGCATGATGACTTT 1440
Qy 1785 GCTGATGACATCTCCCTGCTCAAGTGAAGAGGCGCCATGGCGAGAGATAGAGATTCCCT 1844
Db 1441 GCTGATGACATCTCCCTGCTCAAGTGAAGAGGCGCCATGGCGAGAGATAGAGATTCCCT 1500
Qy 1845 GGACCACACCTCCGTGGTTCATTTGGTCAACAAGTAGGAGACACAGATGGCACCTGTGGC 1904
Db 1501 GGACCACACCTCCGTGGTTCATTTGGTCAACAAGTAGGAGACACAGATGGCACCTGTGGC 1560
Qy 1905 CAGAGCACCCTCAGACCCCTCCCAACCCACCAATGCTCTGCTTGTGAGAGAGGAAAA 1964
Db 1561 CAGAGCACCCTCAGACCCCTCCCAACCCACCAATGCTCTGCTTGTGAGAGAGGAAAA 1620
Qy 1965 GCCTGGCAAGTGGTTCAGGGACTGTACCTGTAGGAAACAGAAAAAGAGAAAGAAAG 2024
Db 1621 GCCTGGCAAGTGGTTCAGGGACTGTACCTGTAGGAAACAGAAAAAGAGAAAGAAAG 1680
Qy 2025 CACTCTGCTGGCGGAATACCTTTGGTCACCTCAAAATTAAGTCGGGAAATTCCTGCT 2084
Db 1681 CACTCTGCTGGCGGAATACCTTTGGTCACCTCAAAATTAAGTCGGGAAATTCCTGCT 1740
Qy 2085 TGAACCTTCAGCCCTGAAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT 2144
Db 1741 TGAACCTTCAGCCCTGAAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT 1800
Qy 2145 CTTCTTTTCTTAGTTTCAAGAGTACTGGCATCACAGCAGGTTACCTTTGGCGTGTGTC 2204
Db 1801 CTTCTTTTCTTAGTTTCAAGAGTACTGGCATCACAGCAGGTTACCTTTGGCGTGTGTC 1860
Qy 2205 TGTGTACCCCTGGCAGAGAAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCACTAGGA 2264

Db 1851 TGTGTACCCCTGGCAGAGAAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCACTAGGA 1920
Qy 2265 GAGGATGCACAGTTCCTATTTGCTTTAGAGACAGGGGACTGTATAAACAAGCCTAACATT 2324
Db 1921 GAGGATGCACAGTTCCTATTTGCTTTAGAGACAGGGGACTGTATAAACAAGCCTAACATT 1980
Qy 2325 GGTGCAAAAGATTGCTCTTTGAATT 2348
Db 1981 GGTGCAAAAGATTGCTCTTTGAATT 2004

Search completed: March 1, 2003, 22:17:24
Job time : 354.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2003, 13:38:25 ; Search time 4002 Seconds
(without alignments)
17074.810 Million cell updates/sec

Title: US-09-723-722a-42
Perfect score: 2348
Sequence: 1 ccattgcggccctcacagc.....caagattgcctcttgatt 2348

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			Length	DB	ID	Description
	Score	Match					
1	2342.8	99.8	2526	9	AF190725	AF190725	Homo sapi
2	2341.2	99.7	5878	9	AF201468	AF201468	Homo sapi
3	2331.8	99.3	5814	9	AB032975	AB032975	Homo sapi
4	2181.2	92.9	5757	6	AX364933	AX364933	Sequence
5	2041.4	86.9	2070	9	AF200343	AF200343	Homo sapi
6	2039.8	86.9	2070	6	AX105385	AX105385	Sequence
7	2039.8	86.9	2070	6	AX378015	AX378015	Sequence
8	2024	86.2	2541	6	AF178469	AF178469	Sequence
9	2024	86.2	2541	6	AX002655	AX002655	Sequence
10	2024	86.2	2541	6	ES0816	ES0816	Aspartate p
11	2004	85.3	3252	6	AX062111	AX062111	Sequence
12	2004	85.3	3252	6	AX063201	AX063201	Sequence
13	2004	85.3	3252	6	AX472368	AX472368	Sequence
14	2004	85.3	3252	9	AF200193	AF200193	Homo sapi
15	1881	80.1	1977	6	AX105387	AX105387	Sequence
16	1881	80.1	1977	6	AX378017	AX378017	Sequence
17	1835.4	78.2	2370	6	AF178470	AF178470	Sequence
18	1835.4	78.2	2370	6	AX002657	AX002657	Sequence
19	1835.4	78.2	2370	6	ES0817	ES0817	Aspartate p
20	1660.6	70.7	3880	10	AF190726	AF190726	Mus muscu
21	1550	66.0	2043	6	AX105389	AX105389	Sequence
22	1550	66.0	2043	10	AF200346	AF200346	Mus muscu
23	1511.2	64.4	2158	6	AX401989	AX401989	Sequence
24	1511.2	64.4	2158	10	AF190727	AF190727	Rattus no
25	1506	64.1	1506	9	AF204943	AF204943	Homo sapi
26	1380	58.8	1465	9	AB050436	AB050436	Homo sapi
27	1355.8	57.7	1362	6	AX105411	AX105411	Sequence
28	1355.8	57.7	1380	6	AX105413	AX105413	Sequence
29	1344.4	57.3	1431	9	AF338816	AF338816	Homo sapi
30	1298	55.3	1506	6	AR202196	AR202196	Sequence
31	1293.8	55.1	1341	6	AX105403	AX105403	Sequence
32	1293.8	55.1	1380	6	AX105405	AX105405	Sequence
33	1270.4	54.1	1302	6	AX105407	AX105407	Sequence
34	1266	53.9	1408	9	AB050437	AB050437	Homo sapi
35	1230.4	52.4	1374	9	AF338817	AF338817	Homo sapi
36	1197.4	51.0	1287	6	AX105432	AX105432	Sequence
37	1197.4	51.0	1305	6	AX105434	AX105434	Sequence
38	1180	50.3	1278	6	AX105409	AX105409	Sequence
39	1116	47.5	1333	9	AB050438	AB050438	Homo sapi
40	990	42.2	1114	9	AF161367	AF161367	Homo sapi
C 41	781.2	33.3	134278	9	AP001822	AP001822	Homo sapi
C 42	781.2	33.3	199892	2	AC020997	AC020997	Homo sapi
C 43	779.6	33.2	98305	2	AP000685	AP000685	Homo sapi
C 44	779.6	33.2	149843	9	AP000892	AP000892	Homo sapi
C 45	779.6	33.2	162610	2	AP000761	AP000761	Homo sapi

ALIGNMENTS

RESULT 1
AF190725
LOCUS
DEFINITION Homo sapiens beta-site APP cleaving enzyme (BACE) mRNA, complete
cvs
ACCESSION AF190725
VERSION AF190725.1 GI:6118538
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2526)
Vassar,R., Bennett,B.D., Babu-Khan,S., Kahn,S., Mendtaz,E.A.,
Denis,P., Teplow,D.B., Ross,S., Amarante,P., Loeloff,R., Luo,Y.,

Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M.A.,
Biere, A.L., Curran, E., Burgess, T., Louis, J.C., Collins, F.,
Trenor, J., Rogers, G. and Citron, M.
Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
the transmembrane aspartic protease BACE

Science 286 (5440), 735-741 (1999)
20002972
10531052

REFERENCE
2 (bases 1 to 2526)
Bennett, B.D., Vassar, R. and Citron, M.

Direct Submission
Submitted (29-SEP-1999) Neuroscience, Amgen Inc., One Amgen Center
Dr., Thousand Oaks, CA 91320-1799, USA

JOURNAL
TITLE
JOURNAL
FEATURES
source

Location/Qualifiers
1. .2526
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene
1. .2526
/gene="BACE"

CDS
454..1959
/gene="BACE"

/codon_start=1
/product="beta-site APP cleaving enzyme"

/protein_id="AA04142.1"

/db_xref="GI:6118539"

/translation="MAQALPWLMLMGAGVLPAGHTGHCIRLRLSLGLGAPLRLP
RETDEPEPPGRGSGFVEMVDNLRGSGGYVEMTGVSPOTLNLIVDTGSSNFVAG
AAPPELHRVYQRLSTYRDLRGVYVPTGKWEGLGTLVSTPHGPNVTVRANI
AATPEDELFINGSNIEGILGLVAETARDPDSLPFPDLSLVKQTHVPLFSLQCGA
GPLNQSEVLASVGGSMIIGIDHSITGSLWYTPIRREWEYEVILRVINEIGDLKM
DCKEYNDKSIDVSGTNNLRPKVFEAAVKSIAKASTKEFPDLMGEQLVCMQAG
TPWNIFPVSILMGEVTVNOSFRITILPOQYLRPEVDATSDQDCYKFAISQSTGT
VMGAVIMEGFYVYFDRARKRIGFAVSACHVDEFRTAAVEGPEVTLDMEDGYNIPOT
DESTLMTIAVYMAAICALEMPLICLMVCOMWCLRLRQHQHDDFADDSILK"

BASE COUNT 534 a 781 c 715 g 456 t

ORIGIN

Query Match 99.8%; Score 2342.8; DB 9; Length 2526;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2344; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCATGCGGCGCCCTCACAGGCGCGCGGGAGCGCCGAGCGCGCTGCGCCAGCGTGGCGCGCG 60
DB 149 CCATGCGGCGCCCTCCACAGCGCGCGGGAGCGCGCGCGCTGCGCCAGCGTGGCGCGCG 208
QY 61 CCGTGCAGATGTAGCGGCGTCCGAGTCCAGCGCTCTCCCGTCTCCCGTCTCTCGCGAT 120
DB 209 CCGTGCAGATGTAGCGGCGTCCGAGTCCAGCGCTCTCCCGTCTCTCCCGTCTCTCGCGAT 268
QY 121 CTCCTCTGACCGCTCTCCAGCGCGCGGACCGCGGGGCTGGCCCGAGCGGCGCTTGACGCGCCT 180
DB 269 CTCCTCTGACCGCTCTCCAGCGCGCGGACCGCGGGGCTGGCCCGAGCGGCGCTTGACGCGCCT 328
QY 181 GCGTCTGATGCCCGCCCAAGCTCTCTCTCTGAGAGCCACAGCGCCACCGAGACTTGG 240
DB 329 GCGTCTGATGCCCGCCCAAGCTCTCTCTCTGAGAGCCACAGCGCCACCGAGACTTGG 388
QY 241 GCGGAGCGCGGAGCGAGCGAGTGGGCGAGTCCGAGCGCCAGAGCGCGCGGAGCGCGGCG 300
DB 389 GCGGAGCGCGGAGCGAGCGAGTGGGCGAGTCCGAGCGCCAGAGCGCGCGGAGCGCGGCG 448
QY 301 CCACCATGGCGCCCAAGCGCTGCGCTTGGCTCTCTCTGAGAGCGCGGCGGCGGAGTCTGCGCTG 360
DB 449 CCACCATGGCGCCCAAGCGCTGCGCTTGGCTCTCTCTGAGAGCGCGGCGGAGTCTGCGCTG 508
QY 361 CCCACGCGCCAGCGAGCGAGTCCGCTGCGCTGCGCGCGCGCGCTGGGGGGCGCGCGCG 420
DB 509 CCCACGCGCCAGCGAGCGAGTCCGCTGCGCTGCGCGCGCGCGCTGGGGGGCGCGCGCG 568
QY 421 TGGGGCTGGGGCTGCGCGCGGAGACCGAGAGCGCGCGGAGCGCGCGCGCGCGGAGGCGCA 480
DB 569 TGGGGCTGGGGCTGCGCGCGGAGACCGAGAGCGCGCGGAGCGCGCGCGCGCGGAGGCGCA 628

QY 481 GCTTTTGGAGATGGTGACAACTGAGGGGCAAGTCTGGGGGAGGCTACTACCTGGAGA 540
DB 629 GCTTTTGGAGATGGTGACAACTGAGGGGCAAGTCTGGGGGAGGCTACTACCTGGAGA 688
QY 541 TGACCGTGGGAGCGCCCGCGAGAGCGCTCAACATCTCTGGTGATACAGGAGCAGTAAC 600
DB 689 TGACCGTGGGAGCGCCCGCGAGAGCGCTCAACATCTCTGGTGATACAGGAGCAGTAAC 748
QY 601 TTCCAGTGGGTGGTGGCGCCCGCACCCCTTCTCTGCATCGCTACTACCGAGGCAAGCTGTCCA 660
DB 749 TTCCAGTGGGTGGTGGCGCCCGCACCCCTTCTCTGCATCGCTACTACCGAGGAGCTGTCCA 808
QY 661 GCACATACCGGAGCTCTCGGAGGCTGTGTATGTGCGCTACACCCAGGCAAGTGGGAAG 720
DB 809 GCACATACCGGAGCTCTCGGAGGCTGTGTATGTGCGCTACACCCAGGCAAGTGGGAAG 868
QY 721 GCGAGCTGGGACCGGAGCTGGTAAGCATCCCGCATGGCCCAACAGCTACTGTGGGTGCCA 780
DB 869 GCGAGCTGGGACCGGAGCTGGTAAGCATCCCGCATGGCCCAACAGCTACTGTGGGTGCCA 928
QY 781 ACATTTGCTGCATCACTGAATCAGACAAAGTTCTTCAATCAAGCGCTCCCACTGGGAAGGCA 840
DB 929 ACATTTGCTGCATCACTGAATCAGACAAAGTTCTTCAATCAAGCGCTCCCACTGGGAAGGCA 988
QY 841 TCCTGGGGCTGGCTATCTGTAGATTTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTTG 900
DB 989 TCCTGGGGCTGGCTATCTGTAGATTTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTTG 1048
QY 901 ACTCTCTGGTAAAGCAGACCCACGTTCCCAAGCTCTTCTCCCTGCAGCTTTGTGGTGGTG 960
DB 1049 ACTCTCTGGTAAAGCAGACCCACGTTCCCAAGCTCTTCTCCCTGCAGCTTTGTGGTGGTG 1108
QY 961 GCTTCCCGCTCAACGAGTCTGAAGTGGTGGCTCTCTCGAGGAGGAGCATGATCATTTGAG 1020
DB 1109 GCTTCCCGCTCAACGAGTCTGAAGTGGTGGCTCTCTCGAGGAGGAGCATGATCATTTGAG 1168
QY 1021 GTATCGACCACTCGCTGTACACAGCAGTCTCTGGTATATACCCATCCCGGGGAGTGGT 1080
DB 1169 GTATCGACCACTCGCTGTACACAGCAGTCTCTGGTATATACCCATCCCGGGGAGTGGT 1228
QY 1081 ATTATGAGTGTATCTTGTGGGTGGAGATCAATGGACAGGATCTGAAAATGACATGCA 1140
DB 1229 ATTATGAGTGTATCTTGTGGGTGGAGATCAATGGACAGGATCTGAAAATGACATGCA 1288
QY 1141 AGGAGTACAACATATGACAAGAGCATTTGGACAGTGGCACCAACCTCTCGTTTGGCGCA 1200
DB 1289 AGGAGTACAACATATGACAAGAGCATTTGGACAGTGGCACCAACCTCTCGTTTGGCGCA 1348
QY 1201 AGAAGTGTGGAGCTGACGTCAAAATCCATCAAGGAGCGCTCTCCACGGAGAGTTC 1260
DB 1349 AGAAGTGTGGAGCTGACGTCAAAATCCATCAAGGAGCGCTCTCCACGGAGAGTTC 1408
QY 1261 CTGATGGTTTCTGGCTAGGAGAGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1320
DB 1409 CTGATGGTTTCTGGCTAGGAGAGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1468
QY 1321 ACATTTTCCAGTCTCACTCTACCTTAATGGGTGAGGTATACCAACAGTTCCTTCCGCA 1380
DB 1469 ACATTTTCCAGTCTCACTCTACCTTAATGGGTGAGGTATACCAACAGTTCCTTCCGCA 1528
QY 1381 TCACCATCTCTTCCGAGCAATACCTGCGGCGCAGTGGAGAGTGGGCCACGTCCTCCAAAG 1440
DB 1529 TCACCATCTCTTCCGAGCAATACCTGCGGCGCAGTGGAGAGTGGGCCACGTCCTCCAAAG 1588
QY 1441 ACTGTTACAAGTGGTGGCATCTACAGTCAATCCAGCGGCACTGTTATGGGAGCTCTTATCA 1500
DB 1589 ACTGTTACAAGTGGTGGCATCTACAGTCAATCCAGCGGCACTGTTATGGGAGCTCTTATCA 1648
QY 1501 TGGAGGCTTCTACGTTGTCTTTGATCGGGCGCGGAGGAGGAGTGGCTTGTCTCTCAGCG 1560
DB 1649 TGGAGGCTTCTACGTTGTCTTTGATCGGGCGCGGAGGAGGAGTGGCTTGTCTCTCAGCG 1708
QY 1561 CTTGGCATGTGCACGATGAGTTTCAGGAGCGGAGCGGTTGGAGGCGGCTTTTGTACCTTTGG 1620

|||||
Db 1709 CTTGCCATGTGCAGATGAGTTTCAGGACGCGGTGGAAGGCCCTTTGTCTACCTTGG 1768
Qy 1621 ACATGGAAGACTGTGGCTACAACTCCACAGACAGATGAGTCAACCTCATGACCATAG 1680
Db 1769 ACATGGAAGACTGTGGCTACAACTCCACAGACAGATGAGTCAACCTCATGACCATAG 1828
Qy 1681 CCTATGTCATGGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCTCATGTGTGTC 1740
Db 1829 CCTATGTCATGGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCTCATGTGTGTC 1888
Qy 1741 AGTGGCCCTGCTCCGCTGCTGCGCCACAGCATGATGACTTTGCTGATGACATCTCCC 1800
Db 1889 AGTGGCCCTGCTCCGCTGCTGCGCCACAGCATGATGACTTTGCTGATGACATCTCCC 1948
Qy 1801 TGCTGAAGTGAAGAGGCCCATGGGCAGAGATAGATTCCTCCCTGGACACACCTCCGCTG 1860
Db 1949 TGCTGAAGTGAAGAGGCCCATGGGCAGAGATAGATTCCTCCCTGGACACACCTCCGCTG 2008
Qy 1861 GTTCACCTTTGGTCAACAAGTAGGACACACAGATGACCTGTGGCCAGAGACCTTCAGGAC 1920
Db 2009 GTTCACCTTTGGTCAACAAGTAGGACACACAGATGACCTGTGGCCAGAGACCTTCAGGAC 2068
Qy 1921 CTTCCCCACCCACCAATGCTCTGCTTGTGAGAGAGGAAAAGCTGGCAAGGTGGT 1980
Db 2069 CTTCCCCACCCACCAATGCTCTGCTTGTGAGAGAGGAAAAGCTGGCAAGGTGGT 2128
Qy 1981 TCCAGGACTGTACCTGTAGGAACACAGAAAGAGAAAGACACTCTGCTGGCGGGA 2040
Db 2129 TCCAGGACTGTACCTGTAGGAACACAGAAAGAGAAAGACACTCTGCTGGCGGGA 2188
Qy 2041 ATACTCTGTGTACCTCAAAATTTAAAGTCGGGAAATTCCTGCTGTTGAAACTTCAGCCCTG 2100
Db 2189 ATACTCTGTGTACCTCAAAATTTAAAGTCGGGAAATTCCTGCTGTTGAAACTTCAGCCCTG 2248
Qy 2101 AACCTTTGTCCACCATCTCTTAAATTTCTCCAAACCCAAAGATTTCTTTTCTTAGTTT 2160
Db 2249 AACCTTTGTCCACCATCTCTTAAATTTCTCCAAACCCAAAGATTTCTTTTCTTAGTTT 2308
Qy 2161 CAGAAGTACTGGCATACACGACGAGTTACCTTGGCGGTGTGCTCCCTGGTACCTTGGCAG 2220
Db 2309 CAGAAGTACTGGCATACACGACGAGTTACCTTGGCGGTGTGCTCCCTGGTACCTTGGCAG 2368
Qy 2221 AGAAGACACCAAGCTTGTTCCTGCTGGCCAAAGTCAAGTAGGAGAGATGACAGATTG 2280
Db 2369 AGAAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAAGTAGGAGAGATGACAGATTG 2428
Qy 2281 CTATTGCTTTAGACAGGAGCTGTATATAACAGCCCTTAACATTTGGTGCAAAAGATTGCT 2340
Db 2429 CTATTGCTTTAGACAGGAGCTGTATATAACAGCCCTTAACATTTGGTGCAAAAGATTGCT 2488
Qy 2341 CTTGCAATT 2348
Db 2489 CTTGCAATT 2496

RESULT 2
AF201468
LOCUS AF201468 5878 bp mRNA linear PRI 19-DEC-1999
DEFINITION Homo sapiens APP beta-secretase mRNA, complete cds.
ACCESSION AF201468
VERSION AF201468.1 GI:6601444
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5878)
Sinha, S., Anderson, J.P., Barbour, R., Basi, G.S., Caccavello, R.,
Davis, D., Doan, M., Dovey, H.F., Frigon, N., Hong, J.,
Jacobson-Croak, K., Jewett, N., Keim, P., Knops, J., Lieberburg, I.,
Power, M., Tan, H., Tatsuno, G., Tung, J., Schenk, D., Seubert, P.,
Suomensaari, S., Wang, S., Walker, D., Zhao, J., McConlogue, L. and

John, V.
Purification and cloning of amyloid precursor protein
beta-secretase from human brain
NATURE (02-06-2000) 405:535-540 (1999)
MEDLINE 20050117
PUBMED 10591214
REFERENCE 2 (bases 1 to 5878)
AUTHORS Basi, G.S., Power, M.D., Wang, S., Tatsuno, G., Frigon, N., Doan, M.,
Hong, G., Keim, P., Anderson, J., Sinha, S. and McConlogue, L.M.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1999) Gene Expression Group, Elan
Pharmaceuticals, Inc., 800 Gateway Blvd., S. San Francisco, CA
94080, USA
FEATURES
Location/Qualifiers
1..5878
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="D11S4352"
/tissue_type="brain"
455..1960
/function="cleaves APP at the beta-cleavage site"
/note="membrane type aspartyl protease"
/codon_start=1
/product="APP beta-secretase"
/protein_id="A6018982.1"
/db_xref="GI:6601445"
/translation="MAQALPWLILWMGAVLPAHGTQHGIRLPLRSLGGAPLGLRLP
REYDEPEEPGRGSEVMYDNLNRKSGQYVEMTVGSPQTLNILLVDTGSSNFVAG
AAPFLPHRYQLSTYRDLRGVYVPTGKWEGLGTLVSVIPGPNVTVRANI
AATLSEDKPFIINGSNMGIAGLAEIARPDLSLEPFDSLKVTHVNPILFSLQCCA
GFLNDSVLAASVGGMIIGIDSLYTLGSLWTPTRREWYEVIIIVRVEINGDLKAM
DCKEYADKSLIVDSGTNLRPKKVEAAVKSIKAASSTKFPDGNLGLQVLCWQAG
TTWNFIPEVLSLYLMGEVNTQSFRTILPOYLRPVEDVATSDQCKYKFAISQSTGT
VMGAVIMEGVYVDFDRARRKIRFASACHVTAAGVFPVTLDMEDCGYNIPOT
DESTLMTIAYVMAAICALEMPLCLMVCQWRCLRLRQQHDDFADDTSLK"
BASE COUNT 1547 a 1518 c 1405 g 1408 t
ORIGIN

Query Match 99.7%; Score 2341.2; DB 9; Length 5878;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2343; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CCATGCGCGCCCTCCACAGCCCGCGGAGCCCGCCGCTGCCCAGCTGCGCCGCG 60
|||||
Db 150 CCATGCGCGCCCTCCACAGCCCGCGGAGCCCGCCGCTGCCCAGCTGCGCCGCG 209
|||||
Qy 61 CSGTGGCGATGAGCGGCTCCGGATCCAGCCTCTCCCTGCTCCCTGCTCTGCGGAT 120
|||||
Db 210 CCGTGGCGATGAGCGGCTCCGGATCCAGCCTCTCCCTGCTCCCTGCTCTGCGGAT 269
|||||
Qy 121 CTCCTCTGACCGCTCTCCACAGCCCGCGGAGCCCGCCGCTGCGCCAGGCCCT 180
|||||
Db 270 CTCCTCTGACCGCTCTCCACAGCCCGCGGAGCCCGCCGCTGCGCCAGGCCCT 329
|||||
Qy 181 GCGTCTCTGATGCCCCCAAGCTCCTCTCTGAGAAGCCACAGCACCACCACTTGG 240
|||||
Db 330 GCGTCTCTGATGCCCCCAAGCTCCTCTCTGAGAAGCCACAGCACCACCACTTGG 389
|||||
Qy 241 GGCAGCGCCGACGAGGACGCTGGGCCAGTGGCAGCCAGAGCCCGCCAGAGCGGGCC 300
|||||
Db 390 GGCAGCGCCGACGAGGACGCTGGGCCAGTGGCAGCCAGAGCCCGCCAGAGCGGGCC 449
|||||
Qy 301 CCACATGGCCCAAGCCTGCCCTGGCTCTCTGCTGATGGCGCGGAGTGTGCTG 360
|||||
Db 450 CCACATGGCCCAAGCCTGCCCTGGCTCTCTGCTGATGGCGCGGAGTGTGCTG 509
|||||
Qy 361 CCCAGCGCACCACAGCAGGCTCCGCTGCTGCGCAGCGCTGCGGCGCGCCGCC 420
|||||
Db 510 CCCAGCGCACCACAGCAGGCTCCGCTGCTGCGCAGCGCTGCGGCGCGCCGCC 569
|||||
Qy 421 TGGGGCTGCGGCTGCGCCCGGAGACCCAGCAAGAGCCCGAGAGCCCGCGGAGGCA 480
|||||

Db 570 TGGGGCTGGGGCTGCCCCGGGAGACCGACGAAGACCGGAGGCCCGCCGGAGGGGCA 629
QY 481 GCTTTTGTGGAGATGTTGGACAACCTGAGGGCAAGTGGGGCAGGCTACTACGTGGAGA 540
Db 630 CTTTGTGGAGATGTTGGACAACCTGAGGGCAAGTGGGGCAGGCTACTACGTGGAGA 689
QY 541 TGACCGTGGGAGCCCCCGGACAGCGTCAACATCTCTGGTGGATACAGGACGAGTAAT 600
Db 690 TGACCGTGGGAGCCCCCGGACAGCGTCAACATCTCTGGTGGATACAGGACGAGTAAT 749
QY 601 TTGCAGTGGGTGCTGCCCCCGGACCGTCTCTGATGCTACTACAGAGGAGCTGTCCA 660
Db 750 TTGCAGTGGGTGCTGCCCCCGGACCGTCTCTGATGCTACTACAGAGGAGCTGTCCA 809
QY 661 GCACATACCGGACCTCCCGAAGGGTGTATGTGCCCCACACCCAGGCAAGTGGGAAG 720
Db 810 GCACATACCGGACCTCCCGAAGGGTGTATGTGCCCCACACCCAGGCAAGTGGGAAG 869
QY 721 GGGAGCTGGGCAACCGACCTGGTAAGCATCCCCCATGGCCCCCAAGTCACTGTGCGTGCCA 780
Db 870 GGGAGCTGGGCAACCGACCTGGTAAGCATCCCCCATGGCCCCCAAGTCACTGTGCGTGCCA 929
QY 781 ACATTTGCTGCCATCACTGAATCAGACAAGTCTTTCATCAACGGCTCCCACTGGGAAGGCA 840
Db 930 ACATTTGCTGCCATCACTGAATCAGACAAGTCTTTCATCAACGGCTCCCACTGGGAAGGCA 989
QY 841 TCCTGGGGCTGGCCTATGCTGAGATTGCCAGGCGCTGACGACTCCCTGGAGCCTTTCTTTG 900
Db 990 TCCTGGGGCTGGCCTATGCTGAGATTGCCAGGCGCTGACGACTCCCTGGAGCCTTTCTTTG 1049
QY 901 ACTCTCTGGTAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTG 960
Db 1050 ACTCTCTGGTAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTG 1109
QY 961 GCTTCCCTCTCAACCACTCTGAAGTGTCTGGCTCTCTCGGAGGAGCATGATCTTGGAG 1020
Db 1110 GCTTCCCTCTCAACCACTCTGAAGTGTCTGGCTCTCTCGGAGGAGCATGATCTTGGAG 1169
QY 1021 GTATCGAACCTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGT 1080
Db 1170 GTATCGAACCTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGT 1229
QY 1081 ATTATGAGGTGATCTTGTGGGGTGGAGATCAATGACAGCATCTGAATATGGACTGCA 1140
Db 1230 ATTATGAGGTGATCTTGTGGGGTGGAGATCAATGACAGCATCTGAATATGGACTGCA 1289
QY 1141 AGGAGTACAACCTATGACAAGAGCATTTGGACAGCTGGCACCAACCACTCTCTTTGCCCA 1200
Db 1290 AGGAGTACAACCTATGACAAGAGCATTTGGACAGCTGGCACCAACCACTCTCTTTGCCCA 1349
QY 1201 AGAAAGTGTGTTGAAGTGCAGTCAAAATCCATCAAGGAGCGCTCTCCACGGAGAGTTCC 1260
Db 1350 AGAAAGTGTGTTGAAGTGCAGTCAAAATCCATCAAGGAGCGCTCTCCACGGAGAGTTCC 1409
QY 1261 CTGATGTTTCTGGCTAGAGAGCAGCTGTGTGCTGGCAAGCAGGACCCCTTGGGA 1320
Db 1410 CTGATGTTTCTGGCTAGAGAGCAGCTGTGTGCTGGCAAGCAGGACCCCTTGGGA 1469
QY 1321 ACATTTTCCAGTCTCTCACTCTACCTAATGGTGGAGTTACCAACCACTCTCTCCGCA 1380
Db 1470 ACATTTTCCAGTCTCTCACTCTACCTAATGGTGGAGTTACCAACCACTCTCTCCGCA 1529
QY 1381 TCACCATCTCTCCGACGAATACCTGGGGCAGTGGAGATGTGGCCACGTCCTCCAGAGCG 1440
Db 1530 TCACCATCTCTCCGACGAATACCTGGGGCAGTGGAGATGTGGCCACGTCCTCCAGAGCG 1589
QY 1441 ACTGTTACAAGTTGGCATCTCAAGTCTACAGTCTCCAGGGGACGTGTATGGAGCTGTATCA 1500
Db 1590 ACTGTTACAAGTTGGCATCTCAAGTCTACAGTCTCCAGGGGACGTGTATGGAGCTGTATCA 1649
QY 1501 TGGAGGCTCTACGTTGCTTTGATCGGGCCGCAAAACGAATTTGGCTTTGCTGCAGCG 1560
Db 1650 TGGAGGCTCTACGTTGCTTTGATCGGGCCGCAAAACGAATTTGGCTTTGCTGCAGCG 1709

QY 1561 CTTGCCATGTGCACGATGAGTTTCAGGACGGCAGCGGTGAAGCCCTTTTGTACACCTTGG 1620
Db 1710 CTTGCCATGTGCACGATGAGTTTCAGGACGGCAGCGGTGAAGCCCTTTTGTACACCTTGG 1769
QY 1621 ACATGGAAGACTGTGGCTAGCAACATTCACAGACAGATGAGTCAACCCCTCATGACCATAG 1680
Db 1770 ACATGGAAGACTGTGGCTAGCAACATTCACAGACAGATGAGTCAACCCCTCATGACCATAG 1829
QY 1681 CCTATGTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTCTGCCCTCATGGTGTGTC 1740
Db 1830 CCTATGTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTCTGCCCTCATGGTGTGTC 1889
QY 1741 AGTGGCGCTCCCTCCGCTGCTGCCAGCAGCATGATGACTTTGCTGATGACATCTCCC 1800
Db 1890 AGTGGCGCTCCCTCCGCTGCTGCCAGCAGCATGATGACTTTGCTGATGACATCTCCC 1949
QY 1801 TGCTGAAGTGAAGAGGCGCCATGGGCAAGAATAGAGATTCCTCTGGACACACCTCCGTG 1860
Db 1950 TGCTGAAGTGAAGAGGCGCCATGGGCAAGAATAGAGATTCCTCTGGACACACCTCCGTG 2009
QY 1861 GTTCACTTTGGTGCACAGTAGGAGACACAGATGCACCTGTGGCCAGACACCTCAGGAC 1920
Db 2010 GTTCACTTTGGTGCACAGTAGGAGACACAGATGGCACCTGTGGCCAGACACCTCAGGAC 2069
QY 1921 CTTCCCAACCCCAATGCTCTGCTTGTGAGAGAGGAAAGCTGGCAAGGTGGGT 1980
Db 2070 CTTCCCAACCCCAATGCTCTGCTTGTGAGAGAGGAAAGCTGGCAAGGTGGGT 2129
QY 1981 TCCAGGAGCTGTACCTGTAGGAAACAGAAAGAGAGAAAGCACTCTGCTGGCGGGA 2040
Db 2130 TCCAGGAGCTGTACCTGTAGGAAACAGAAAGAGAGAAAGCACTCTGCTGGCGGGA 2189
QY 2041 ATACTCTTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTGCTTCAAACTTTCAGCCCTG 2100
Db 2190 ATACTCTTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTGCTTCAAACTTTCAGCCCTG 2249
QY 2101 AACCTTTGTCACCATCTCTTAAATTCCTCAACCCAAAGATATCTCTTTTCTTACTT 2160
Db 2250 AACCTTTGTCACCATCTCTTAAATTCCTCAACCCAAAGATATCTCTTTTCTTACTT 2309
QY 2161 CAGAAGTACTGGCATCACAGCAGGTTACCTTGGCGGTGTCCTTGTGTACCTCCGCGAG 2220
Db 2310 CAGAAGTACTGGCATCACAGCAGGTTACCTTGGCGGTGTCCTTGTGTACCTCCGCGAG 2369
QY 2221 AGAAGAGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTG 2280
Db 2370 AGAAGAGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTG 2429
QY 2281 CTATTTGCTTTAGACACAGGAGCTGTATAACACAGCCCTAACATTTGGTGCAGAGATTGCCCT 2340
Db 2430 CTATTTGCTTTAGACACAGGAGCTGTATAACACAGCCCTAACATTTGGTGCAGAGATTGCCCT 2489
QY 2341 CTTGAATT 2348
Db 2490 CTTGAATT 2497

RESULT 3
AB032975
LOCUS
DEFINITION Homo sapiens mRNA for KIAA1149 protein, partial cds.
ACCESSION AB032975
VERSION AB032975.2 GI:14133242
KEYWORDS
SOURCE Homo sapiens adult male brain cDNA to mRNA, clone_lib:pBluescriptII
SK plus clone:fg04087.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Hiroseawa,M., Nagase,T., Ishikawa,K., Kikuno,R., Nomura,N. and Ohara,O.

TITLE	Characterization of cDNA clones selected by the Genewark analysis
JOURNAL	from-size-fractionated cDNA libraries from human brain
MEDLINE	DNARES. 6 (5): 329-336 (1999)
REFERENCE	20039618
AUTHORS	2 (bases 1 to 5814)
TITLE	Chara,O., Nagase,T. and Kikuno,R.
JOURNAL	Direct Submission
COMMENT	Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913, Fax:+81-438-52-3914) On May 17, 2001 this sequence version replaced gi:63300044.

Query Match	99.3%	Score	2331.8;	DB	9;	Length	5814;
Best Local Similarity	99.8%;	Pred.	No. 0;				
Matches	2344;	Conservative	1;	Mismatches	3;	Indels	1;
Gaps	1;						
QY	1	CCATCGCGGCGCTCACAGCCCGCGGGAGCCGAGCGCCGCTGCCAGGCTGGCGCGCG	60				
Db	138	CCATGCGCGCGCTCCAGCGCCCGGGAGCCGCGCGCTGCCAGGCTGGCGCGCG	197				
QY	61	CSGTGCGGATGTAGCGGGCTCCGGATCCAGCCTCTCCCTGCTCCCTGCTCTGCGGAT	120				
Db	198	CCGTGCGCGATGTAGCGGGCTCCGGATCCAGCCTCTCCCTGCTCCCTGCTCTGCGGAT	257				
QY	121	CTCCCTGACCGCTCTCCAGAGCCCGAGCCGGGGGTGGCCAGGGCCCTGCAGGCCCT	180				
Db	258	CTCCCTGACCGCTCTCCAGAGCCCGAGCCGGGGGTGGCCAGGGCCCTGCAGGCCCT	317				
QY	181	GGCTCTCTGATGCCCCCAAGCTCCTCTCTGAAAGCCACAGCACCCAGACTTGG	240				
Db	318	GGCTCTCTGATGCCCCCAAGCTCCTCTCTGAAAGCCACAGCACCCAGACTTGG	377				
QY	241	GGGAGGCGCCAGGAGCGAGGTGGGCGAGTGCAGGCCAGAGGGCCCGAAGCGCGGGC	300				
Db	378	GGGAGGCGCCAGGAGCGAGGTGGGCGAGTGCAGGCCAGAGGGCCCGAAGCGCGGGC	437				
QY	301	CCACCATGGGCCAAGCCCTGCCTCGCTCTGAGTGGGGCGGGAGTGTGCTCTG	360				

QY 1440 GACTGTTACAAGTTTGGCCATCTCACAGTCTATCCAGGGCACTGTATATGGAGCTGTATC 1499
Db 1578 GACTGTTACAAGTTTGGCCATCTCACAGTCTATCCAGGGCACTGTATATGGAGCTGTATC 1637
QY 1500 ATGGAGGGCTTCTACCTGTCTTGTATCGGGCCGAAACCAATTTGGCTTCTGTCTCAGC 1559
Db 1638 ATGGAGGGCTTCTACCTGTCTTGTATCGGGCCGAAACCAATTTGGCTTCTGTCTCAGC 1697
QY 1560 GCTTGCCATGTGCACGATGAGTTTCAGGACGCGACGGTGGAGGCGCTTTTGTCACTTG 1619
Db 1698 GCTTGCCATGTGCACGATGAGTTTCAGGACGCGACGGTGGAGGCGCTTTTGTCACTTG 1757
QY 1620 GACATGGAAGACTGTGGCTACACATTTCCACAGACAGATGATCAACCTCATGACCATTA 1679
Db 1758 GACATGGAAGACTGTGGCTACACATTTCCACAGACAGATGATCAACCTCATGACCATTA 1817
QY 1680 GCTATGTCATGCTGCCATCTCGCCCTCTTCACTGTGCGCCACTCTGCTCATGCTGTGT 1739
Db 1818 GCTATGTCATGCTGCCATCTCGCCCTCTTCACTGTGCGCCACTCTGCTCATGCTGTGT 1877
QY 1740 CAGTGGCGTGCCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTGTGATGACATCTCC 1799
Db 1878 CAGTGGCGTGCCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTGTGATGACATCTCC 1937
QY 1800 CTGCTGAAGTAGGAGGCCCATGGGAGAGATAGAGATTCCCTGGAGCACACCTCCGT 1859
Db 1938 CTGCTGAAGTAGGAGGCCCATGGGAGAGATAGAGATTCCCTGGAGCACACCTCCGT 1997
QY 1860 GCTTTCACCTTTGCTCACAAGTAGGAGACACAGATGCGCACTGTGGCCAGACGACCTCAGCA 1919
Db 1998 GCTTTCACCTTTGCTCACAAGTAGGAGACACAGATGCGCACTGTGGCCAGACGACCTCAGCA 2057
QY 1920 CCTCCCAACCCACCAATGCTCTGCTTGTATGAGAGAAAGAGTGGCAAGGTGGG 1979
Db 2058 CCTCCCAACCCACCAATGCTCTGCTTGTATGAGAGAAAGAGTGGCAAGGTGGG 2117
QY 1980 TTCCAGGACTGTACCTGTAGGAAACAGAAAGAGAAAGAGTGTGCTGGCGGG 2039
Db 2118 TTCCAGGACTGTACCTGTAGGAAACAGAAAGAGAAAGAGTGTGCTGGCGGG 2177
QY 2040 AATACCTTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTTGAACCTTCAGCCCT 2099
Db 2178 AATACCTTTGGTCACTCAAAATTTAAGTCGGGAAATTTCTGCTTGAACCTTCAGCCCT 2237
QY 2100 GAACCTTTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGTATTTCTTTTCTTAGTT 2159
Db 2238 GAACCTTTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGTATTTCTTTTCTTAGTT 2297
QY 2160 TCAGAACTACTGGCATCACACGAGTTACCTTGGCGTGTGCTCCTGTGGTACCTGGCA 2219
Db 2298 TCAGAACTACTGGCATCACACGAGTTACCTTGGCGTGTGCTCCTGTGGTACCTGGCA 2357
QY 2220 GAGAGAGACCAAGCTTGTTCCTGCTGCGCCAAAGTCAGTAGGAGAGGATGACAGTTT 2279
Db 2358 GAGAGAGACCAAGCTTGTTCCTGCTGCGCCAAAGTCAGTAGGAGAGGATGACAGTTT 2417
QY 2280 GCTATTTGCTTTAGACAGGGACTGTATAAACAAGCTTAACATTTGGTGCAAGATTGCC 2339
Db 2418 GCTATTTGCTTTAGACAGGGACTGTATAAACAAGCTTAACATTTGGTGCAAGATTGCC 2477
QY 2340 TCTTGAAT 2348
Db 2478 TCTTGAAT 2486

RESULT 4
AX364933
LOCUS AX364933 5757 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 84 from Patent WO0206315.
ACCESSION AX364933
VERSION AX364933.1 GI:18696823
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Mintz, L., Freilich, S. and Bernstein, J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 84 24-JAN-2002;
Compugen Ltd. (IL)
FEATURES
source
1..5757
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1485 a 1490 c 1392 g 1388 t 2 others
ORIGIN
Query Match 92.9%; Score 2181.2; DB 6; Length 5757;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 2268; Conservative 1; Mismatches 4; Indels 75; Gaps 1;
QY 1 CCATGCGCGCCCTCACAGCCCGCGGAGCCGCGGCTGCCAGCTGCCAGCTGGCGCGCC 60
Db 150 CCATGCGCGCCCTCCAGCCCGCGGAGCCGCGGCTGCCAGCTGGCGCGCC 209
QY 61 CSGTCCGATGTAGCGGGTCCGGATCCAGCCTCTCCCTGCTTCCGCTGCTTCCGGAT 120
Db 210 CCGTCCGATGTAGCGGGTCCGGATCCAGCCTCTCCCTGCTTCCGCTGCTTCCGGAT 269
QY 121 CTCCTCTGACCGCTCTCCAGCCCGCGGAGCCGCGGCTGGCCAGGCGCTTCAGGCGCT 180
Db 270 CTCCTCTGACCGCTCTCCAGCCCGCGGAGCCGCGGCTGGCCAGGCGCTTCAGGCGCT 329
QY 181 GCGCTCTGATGCGCCCAAGCTCCTCTCTGAGAAGCCAGCAGCACCACCCAGACTTGG 240
Db 330 GCGCTCTGATGCGCCCAAGCTCCTCTCTGAGAAGCCAGCAGCACCACCCAGACTTGG 389
QY 241 GGGCAGGCGCCAGGACGAGTGGCCAGTGGCCAGAGCCAGAGGCGCCGAGGCGCGGCG 300
Db 390 GGGCAGGCGCCAGGACGAGTGGCCAGTGGCCAGAGCCAGAGGCGCCGAGGCGCGGCG 449
QY 301 CCACATGCGCCCAAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 450 CCACATGCGCCCAAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509
QY 361 CCCAGGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 420
Db 510 CCCAGGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 569
QY 421 TGGGCTGGGCTGCGCCCGGAGACGACGAGAGCCGAGAGCCGAGAGCCGAGAGCCGAGAG 480
Db 570 TGGGCTGGGCTGCGCCCGGAGACGACGAGAGCCGAGAGCCGAGAGCCGAGAGCCGAGAG 629
QY 481 GCTTTGTGAGATGTTGAGAACCTTGAGGGCAAGTGGGGCAAGTGGGGCAAGTGGGGCA 540
Db 630 GCTTTGTGAGATGTTGAGAACCTTGAGGGCAAGTGGGGCAAGTGGGGCAAGTGGGGCA 689
QY 541 TCACCGTGGGACCGCCCGCGCAGACGCTCAACATCTCTGCTGATACAGGACGAGTAACT 600
Db 690 TGACCGTGGGACCGCCCGCGCAGACGCTCAACATCTCTGCTGATACAGGACGAGTAACT 749
QY 601 TTGAGTGGGCTGCTGCGCCCGCAGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 750 TTGAGTGGGCTGCTGCGCCCGCAGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
QY 661 GCACATACCGGACCTTCCGGAGGGTGTGTATGTGCTTACACCCAGGCGAAGTGGGAAG 720
Db 810 GCACATACCGGACCTTCCGGAGGGTGTGTATGTGCTTACACCCAGGCGAAGTGGGAAG 869
QY 721 GGGAGCTGGGACCGACCTTGGTAAGCATCCCGCTGCGCCCAAGCTTCTGCTGCTGCTGCTG 780
Db 870 GGGAGCTGGGACCGACCTTGGTAAGCATCCCGCTGCGCCCAAGCTTCTGCTGCTGCTGCTG 929
QY 781 ACATTTGCTGCCATCACTGAATCAGACAAGTTCTTTCATCAAGGCTTCCAACTGGGAAGCA 840

Db 930 ACATTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCCACTGGGAAGCA 989
Qy 841 TCCTGGGGCTGGCCATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTCTTTG 900
Db 990 TCCTGGGGCTGGCCATGCTGAGATTGCCAG-----1020
Qy 901 ACTCTCTGGTAAAGCAGACACCCAGCTTCCCAACCTCTTCTCCCTGCAGCTTTTGGTCTG 960
Db 1021 -----CCTTTGGTCTG 1034
Qy 961 GCTTCCCTCAACCAAGTCTGAAGTGTGCCCTCTGTGCGAGGAGCATGATCATTTGGAG 1020
Db 1035 GCTTCCCTCAACCAAGTCTGAAGTGTGCCCTCTGTGCGAGGAGCATGATCATTTGGAG 1094
Qy 1021 GTATCGACCACTCGCTGTACACAGGAGCTCTGTGATATACCCATCCGCGGGAGTGGT 1080
Db 1095 GTATCGACCACTCGCTGTACACAGGAGCTCTGTGATATACCCATCCGCGGGAGTGGT 1154
Qy 1081 ATTATGAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCA 1140
Db 1155 ATTATGAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCA 1214
Qy 1141 AGGAGTACAATATGACAAGAGCATTTGGACAGTGGCACCAACCTTTCGTTTGGCCA 1200
Db 1215 AGGAGTACAATATGACAAGAGCATTTGGACAGTGGCACCAACCTTTCGTTTGGCCA 1274
Qy 1201 AGAAGTGTGTTGAGCTGCAGTCAATTCATCAAGGAGGCTCTCCACGGAGAGTTCC 1260
Db 1275 AGAAGTGTGTTGAGCTGCAGTCAATTCATCAAGGAGGCTCTCCACGGAGAGTTCC 1334
Qy 1261 CTGATGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCAACCTTGGGA 1320
Db 1335 CTGATGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCAACCTTGGGA 1394
Qy 1321 ACATTTCCAGTCACTCACTTACCTAATGGGTGAGTTACCAACAGTCCCTTCGCGCA 1380
Db 1395 ACATTTCCAGTCACTCACTTACCTAATGGGTGAGTTACCAACAGTCCCTTCGCGCA 1454
Qy 1381 TCACCATCTTCCGACGAATACCTCGGCGCAGTGGAGATGTGGCCACGCTCCCAAGC 1440
Db 1455 TCACCATCTTCCGACGAATACCTCGGCGCAGTGGAGATGTGGCCACGCTCCCAAGC 1514
Qy 1441 ACTGTTACAAGTTTGCCATCTCAGATCATCCAGCGCACTGTTATGGGAGCTGTTATCA 1500
Db 1515 ACTGTTACAAGTTTGCCATCTCAGATCATCCAGCGCACTGTTATGGAGCTGTTATCA 1574
Qy 1501 TGGAGGCTCTACGTTGCTTTGATCGGCGCCGAAACGAATTTGCTGCTCAGCG 1560
Db 1575 TGGAGGCTCTACGTTGCTTTGATCGGCGCCGAAACGAATTTGCTGCTCAGCG 1634
Qy 1561 CTTGCCATGTGCAGGATGAGTTTCAGGAGCGCAGCGGTGGAGGCGCTTTTCTCACCTGG 1620
Db 1635 CTTGCCATGTGCAGGATGAGTTTCAGGAGCGCAGCGGTGGAGGCGCTTTTCTCACCTGG 1694
Qy 1621 ACATGGAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCTCATGACCATAG 1680
Db 1695 ACATGGAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCTCATGACCATAG 1754
Qy 1681 CCTATGATGCTGCTGATCGGCGCTTTCATGCTGCGCACTGCTGCTCATGCTGCTGTC 1740
Db 1755 CCTATGATGCTGCTGATCGGCGCTTTCATGCTGCGCACTGCTGCTCATGCTGCTGTC 1814
Qy 1741 AGTGGCGCTGCCCTGCTGCTGCGCAGCAGCATGATGATTTGCTGATGACATCTCCC 1800
Db 1815 AGTGGCGCTGCCCTGCTGCTGCGCAGCAGCATGATGATTTGCTGATGACATCTCCC 1874
Qy 1801 TGTGTAAGTGAAGGAGGCGGCTTGGCAGAAGATAGAGATTTCCCTGGACCAACCTCCGCTG 1860
Db 1875 TGTGTAAGTGAAGGAGGCGGCTTGGCAGAAGATAGAGATTTCCCTGGACCAACCTCCGCTG 1934
Qy 1861 GTTCACCTTTGCTACAAGTAGGAGACACAGATGTCACCTGTGGCCAGAGCACCTCAGGAC 1920
|||||

Db 1935 GTTCACCTTTGGTCAACAAGTAGGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGAC 1994
Qy 1921 CTTCCCCACCCACCAATGCTCTGCTTGTGAGAGGAAAGGCTGGCAAGTGGGT 1980
Db 1995 CTTCCCCACCCACCAATGCTCTGCTTGTGAGAGGAAAGGCTGGCAAGTGGGT 2054
Qy 1981 TCCAGGAGCTGTACCTGTAGGAAACAGAAAGAGAGAAAGCACTCTGCTGGCGGA 2040
Db 2055 TCCAGGAGCTGTACCTGTAGGAAACAGAAAGAGAGAAAGCACTCTGCTGGCGGA 2114
Qy 2041 ATACTTTGGTACCTCAAAATTTAAAGTCGGGAAATTCGTGCTTGAACCTTCAGCCCTG 2100
Db 2115 ATACTTTGGTACCTCAAAATTTAAAGTCGGGAAATTCGTGCTTGAACCTTCAGCCCTG 2174
Qy 2101 AACCTTTGTCACCATCTTAAATTTCTCAACCCAAAGTATTCTTTTCTTAGTTT 2160
Db 2175 AACCTTTGTCACCATCTTAAATTTCTCAACCCAAAGTATTCTTTTCTTAGTTT 2234
Qy 2161 CAGAAGTACTGGCATCACACGCGAGTTTACCTTGGCGTGTGTCCTGTGGTACCCCTGGCAG 2220
Db 2235 CAGAAGTACTGGCATCACACGCGAGTTTACCTTGGCGTGTGTCCTGTGGTACCCCTGGCAG 2294
Qy 2221 AGAAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG 2280
Db 2295 AGAAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG 2354
Qy 2281 CTATTTCTTTAGAGACGAGGACTGTATAACAGGCTTAACATTTGCTTCAAGATTCCT 2340
Db 2355 CTATTTCTTTAGAGACGAGGACTGTATAACAGGCTTAACATTTGCTTCAAGATTCCT 2414
Qy 2341 CTTGAATT 2348
Db 2415 CTTGAATT 2422

RESULT 5

AF200343 2070 bp mRNA linear PRI 12-DEC-1999
LOCUS Homo sapiens chromosome 11 aspartyl protease 2 mRNA, complete cds.
DEFINITION
ACCESSION AF200343
VERSION AF200343.1 GI:6561813

KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 2070).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS
Van, R., Bienkowski, M.J., Shuck, M.E., Miao, H., Tory, M.C.,
Paulley, A.M., Brashier, R., Stratman, N.C., Mathews, W.R., Buhl, A.E.,
Carter, D.B., Tomasselli, A.G., Parodi, L.A., Heinrikson, R.L. and
Guadagno, M.E.

ABSTRACT
Mammalian aspartyl protease with Alzheimer's disease
JOURNAL
NATURE 402, 067015-533-537 (1999)
MEDLINE
20051470

REFERENCE
2 (bases 1 to 2070)
Blenkowski, M.J., Shuck, M.E., Slightom, J.L. and Drong, R.F.
AUTHORS
Direct Submission
TITLE
JOURNAL
Submitted (29-OCT-1999) Genomics Research, PharmaciasUpjohn, 301
Henrietta, Kalamazoo, MI 49007, USA

FEATURES
Location/Qualifiers
1. 2070
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q23.3-24.1"
1. 1506
/codon_start=1
/product="aspartyl protease 2"
/protein_id="AAFI079.1"
/db_xref="GI:6561814"
/translation="MAQALPWLIIWMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLP

CDS
1. 1506
/codon_start=1
/product="aspartyl protease 2"
/protein_id="AAFI079.1"
/db_xref="GI:6561814"
/translation="MAQALPWLIIWMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLP

RETDEPEPEGRGRGSEVEMVNLGRKSGOGYYVEMTVGSPPTUNILVDTGSSNFAVG
AAPHPFLHRYTQRLSSTYRDLRKGVYVYTOGKEGELGTDLSVIPHGNVTVRANI
AAITPESDKFYNGSNWEGILGLAYAIARPDDSLVFOKTHVPLNLSLOLCCA
GFLPNQSVLASVDSCTMLIGDHSLYTGSWLPPIRREWYEVIIIVRVEINGODLKM
DCKEYNDKSLVDSCTMLRPLPKVFEAAVKSJKAASSTEFKPDGFWLGEOLVCHQAG
TTPWNIFPISLYLGEVNTQSFRTIILFOQYLRVEDVATSDQDCFKFALSQSSTGT
VMGAVIMEGFYVFDRAARRKIGFAVSACHVHDEBFTAAVGEFPFVTLDMEDCGYNIPQT
DESTLMTIAYVMAATCALFMLPLCLMVCWMRCRLRQHQHDFADDSLLK*
BASE COUNT 476 a 582 c 563 g 449 t
ORIGIN

Query Match		86.9%;	Score 2041.4;	DB 9;	Length 2070;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2042;		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;
QY	306	ATGGCCCAAGCCCTCCCTGGCTGCTGCTGGATGGCGGGGAGTGCCTCCCTGCCAC	365		
DB	1	ATGGCCCAAGCCCTCCCTGGCTGCTGCTGGATGGCGGGGAGTGCCTCCCTGCCAC	60		
QY	366	GGCACCCAGCAGCGCATCGGCTGCCCTCGCGAGCGCGCTGGGGGGCGCCCTGGGG	425		
DB	61	GGCACCCAGCAGCGCATCGGCTGCCCTCGCGAGCGCGCTGGGGGGCGCCCTGGGG	120		
QY	426	CTGGCGCTGCCCGGAGACCGAGAACGAGCCGAGAGCGCCGCGGAGGGGAGCTTT	485		
DB	121	CTGGCGCTGCCCGGAGACCGAGAACGAGCCGAGAGCGCCGCGGAGGGGAGCTTT	180		
QY	486	GTGGAGATGTTGGACAACCTGAGGGCAAGTGGGGGAGGCTACTAGTGGAGATGACC	545		
DB	181	GTGGAGATGTTGGACAACCTGAGGGCAAGTGGGGGAGGCTACTAGTGGAGATGACC	240		
QY	546	GTGGGACGCCGCCGAGACGCTCAACATCCTGTGTGATACAGGACGAGTAACCTTGCA	605		
DB	241	GTGGGACGCCGCCGAGACGCTCAACATCCTGTGTGATACAGGACGAGTAACCTTGCA	300		
QY	606	GTGGGTGTGCCGCCACCCCTTCCTGCACTGCTACTACAGAGGCGAGCTGTCCAGCACA	665		
DB	301	GTGGGTGTGCCGCCACCCCTTCCTGCACTGCTACTACAGAGGCGAGCTGTCCAGCACA	360		
QY	666	TACCGGGACCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGAAGTGGGAAGGGAG	725		
DB	361	TACCGGGACCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGAAGTGGGAAGGGAG	420		
QY	726	CTGGGCAACCGACCTGGTAAGCATCCCCATGGCCCCCAACCTGCTGCTGCCAACATTT	785		
DB	421	CTGGGCAACCGACCTGGTAAGCATCCCCATGGCCCCCAACCTGCTGCTGCCAACATTT	480		
QY	786	GCTGCCATCACTGANTCAGACAAGTTCTTTCATCAACGGCTCCAACTGGGAGGCACTCTG	845		
DB	481	GCTGCCATCACTGANTCAGACAAGTTCTTTCATCAACGGCTCCAACTGGGAGGCACTCTG	540		
QY	846	GGGCTGGCTATGCTGAGATGCCAGGCTGACGACTCCTCGAGGCTTCTTCTTGACTCT	905		
DB	541	GGGCTGGCTATGCTGAGATGCCAGGCTGACGACTCCTCGAGGCTTCTTCTTGACTCT	600		
QY	906	CTGTTAAAGCAGACCCAGCTTCCCAACCTTCTTCCCTGCGAGCTTTGTGCTGGCTTC	965		
DB	601	CTGTTAAAGCAGACCCAGCTTCCCAACCTTCTTCCCTGCGAGCTTTGTGCTGGCTTC	660		
QY	966	CCCTCAACCAAGTCTGAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC	1025		
DB	661	CCCTCAACCAAGTCTGAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC	720		
QY	1026	GACCACCTGCCTGTACACAGGAGCTCTCTGTATACACCCATCCGGCGGAGTGTATTAT	1085		
DB	721	GACCACCTGCCTGTACACAGGAGCTCTCTGTATACACCCATCCGGCGGAGTGTATTAT	780		
QY	1086	GAGTGTATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	1145		
DB	781	GAGTGTATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	840		
QY	1146	TACAACCTATGACAAGAGCATTTGGACAGTGGGACCAACCAACCTTCGTTGCCCAAGAAA	1205		

DB	841	TACAACTATGACAAGAGCATTTGGAGACATGGGACACCAACCAACCTTCGTTTGCCCAAGAA	900		
QY	1206	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCTCTCCACGAGAAAGTTCCCTGAT	1265		
DB	901	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCTCTCCACGAGAAAGTTCCCTGAT	960		
QY	1266	GGTTTCTGGCTAGGAGAGCAGCTGGTGTGTGGCAAGCAGCACACCCCTTGGAAATTT	1325		
DB	961	GGTTTCTGGCTAGGAGAGCAGCTGGTGTGTGGCAAGCAGCACACCCCTTGGAAATTT	1020		
QY	1326	TTCCAGTCATCTCACTTACCTTAATGGGTGAGGTTACCAACCAACCTTCCTCCGCATACC	1385		
DB	1021	TTCCAGTCATCTCACTTACCTTAATGGGTGAGGTTACCAACCAACCTTCCTCCGCATACC	1080		
QY	1386	ATCCTTCCGACCAATACCTGCGGCCAGTGGCAAGATGTGGCACCTCCCAACAGCAGCTGT	1445		
DB	1081	ATCCTTCCGACCAATACCTGCGGCCAGTGGCAAGATGTGGCACCTCCCAACAGCAGCTGT	1140		
QY	1446	TACAAGTTTGGCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG	1505		
DB	1141	TACAAGTTTGGCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG	1200		
QY	1506	GGCTTCTAGCTTGTCTTTGATCGGGGCCGAAACGAATTTGGCTTTGCTGTCAGCGCTTGC	1565		
DB	1201	GGCTTCTAGCTTGTCTTTGATCGGGGCCGAAACGAATTTGGCTTTGCTGTCAGCGCTTGC	1260		
QY	1566	CATGTGCAGGATGAGTTTCAAGGCGGAGCGTGGAAAGCCCTTTTGTCACTTTGGACATG	1625		
DB	1261	CATGTGCAGGATGAGTTTCAAGGCGGAGCGTGGAAAGCCCTTTTGTCACTTTGGACATG	1320		
QY	1626	GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGGCTTAT	1685		
DB	1321	GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGGCTTAT	1380		
QY	1686	GTCATGGCTGCCATCTGCGCCCTTCTTCACTGCTGCCACTCTGCTCATGCTGCTGCTGCTG	1745		
DB	1381	GTCATGGCTGCCATCTGCGCCCTTCTTCACTGCTGCCACTCTGCTCATGCTGCTGCTGCTG	1440		
QY	1746	CGCTGCCCTCCGCTGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG	1805		
DB	1441	CGCTGCCCTCCGCTGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG	1500		
QY	1806	AAGTGAGGAGGCCCTTGGCAGAGATAGAGATTTCCCTTGACACCACTCCCTGGTGTCA	1865		
DB	1501	AAGTGAGGAGGCCCTTGGCAGAGATAGAGATTTCCCTTGACACCACTCCCTGGTGTCA	1560		
QY	1866	CTTTGGTCACAAGTAGGAGACACAGATGGCAACCTGTGGCCAGAGCACCTTCAGGACCTCC	1925		
DB	1561	CTTTGGTCACAAGTAGGAGACACAGATGGCAACCTGTGGCCAGAGCACCTTCAGGACCTCC	1620		
QY	1926	CCACCCCAAAATGCCCTCTGCTTGTATGGAGAGGAAAGCTTGGCAAGTGGGTTTCCAG	1985		
DB	1621	CCACCCCAAAATGCCCTCTGCTTGTATGGAGAGGAAAGCTTGGCAAGTGGGTTTCCAG	1680		
QY	1986	GGACTGTACCTCTAGGAAACAGAAAGAGAGAAAGAAAGCACTCTGCTGGCGGGAATACT	2045		
DB	1681	GGACTGTACCTCTAGGAAACAGAAAGAGAGAAAGAAAGCACTCTGCTGGCGGGAATACT	1740		
QY	2046	CTTGGTCACCTCAAAATTTAAGTCGGGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2105		
DB	1741	CTTGGTCACCTCAAAATTTAAGTCGGGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800		
QY	2106	TTGTCCACCATTTCTTAAATTTCTCAACCCCAAAAGTATTTCTTTTCTTCTTCTTCTTCTT	2165		
DB	1801	TTGTCCACCATTTCTTAAATTTCTCAACCCCAAAAGTATTTCTTTTCTTCTTCTTCTTCTT	1860		
QY	2166	GTACTGGCATCATCACGAGGTTTACCTTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	2225		
DB	1861	GTACTGGCATCATCACGAGGTTTACCTTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1920		
QY	2226	AGACCAAGCTTCTTCCCTGCTGCGCCAAAGTCACTAGAGAGGATGCACAGTTTGTCTATTT	2285		
DB	1921	AGACCAAGCTTCTTCCCTGCTGCGCCAAAGTCACTAGAGAGGATGCACAGTTTGTCTATTT	1980		


```
QY 1926 CCACCCACCAATGCTCTGCTTGTGATGAGAGGAAAGGCTGGCAAGGTGGGTTCCAG 1985
Db 1621 CCACCCACCAATGCTCTGCTTGTGATGAGAGGAAAGGCTGGCAAGGTGGGTTCCAG 1680
QY 1986 GGACTGTACCTGTAGGAACAGAGAAAGAGAGAAAGAGCACTCTGCTGGCGGGGAATACT 2045
Db 1681 GGACTGTACCTGTAGGAACAGAGAAAGAGAGAAAGAGCACTCTGCTGGCGGGGAATACT 1740
QY 2046 CTGGTACCTCAAAATTTAAGTCGGGAAATTTGCTGCTTCAAACTTTCAGCCCTGAACCT 2105
Db 1741 CTGGTACCTCAAAATTTAAGTCGGGAAATTTGCTGCTTCAAACTTTCAGCCCTGAACCT 1800
QY 2106 TTGTCCACCAATTCCTTTAAATTTCAACCAAGATTTCTTCTTTAGTTTTCAGAA 2165
Db 1801 TTGTCCACCAATTCCTTTAAATTTCAACCAAGATTTCTTCTTTAGTTTTCAGAA 1860
QY 2166 GTACTGGGATCACAGCAGGTACCTTGGCGTGTGCTCCCTGTGTACCCCTGGCAGAGAG 2225
Db 1861 GTACTGGGATCACAGCAGGTACCTTGGCGTGTGCTCCCTGTGTACCCCTGGCAGAGAG 1920
QY 2226 AGACCAAGCTTCTTCCCTGCTGGCCAAAGTCAGTAGGAGGATGCACAGTTTCTATT 2285
Db 1921 AGACCAAGCTTCTTCCCTGCTGGCCAAAGTCAGTAGGAGGATGCACAGTTTCTATT 1980
QY 2286 TGCTTTAGACAGGAGCTGTATAAACAAGCCTAAACATTTGGTCAAAAGATTGCCCTTTGA 2345
Db 1981 TGCTTTAGACAGGAGCTGTATAAACAAGCCTAAACATTTGGTCAAAAGATTGCCCTTTGA 2040
QY 2346 ATT 2348
Db 2041 ATT 2043

RESULT 7
AX378015
LOCUS AX378015 2070 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 1 from Patent WO0206306.
ACCESSION AX378015
VERSION AX378015.1 GI:19574050
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yan, R., Tomasselli, A. G., Gurney, M. E., Emmons, T. L., Bienkowski, M. J.
and Heinrikson, R. L.
TITLE Substrates and assays for _g(b)-secretase activity
JOURNAL Patent: WO 0206306-A 1 24-JAN-2002;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source
1. .2070
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 476 a 583 c 562 g 449 t
ORIGIN
Query Match 86.9%; Score 2039.8; DB 6; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 306 ATGCCCAAGCCCTGCCCTGGCTCTGCTGTGGATGGCGGGGAGTGCTGCTGCCCCAC 365
Db 1 ATGCCCAAGCCCTGCCCTGGCTCTGCTGTGGATGGCGGGGAGTGCTGCTGCCCCAC 60
QY 366 GGCACCCAGCAGGATCCGGCTGCCCTCCGCGAGCGGCTGGGGGGGGGGGGGGGGGG 425
Db 61 GGCACCCAGCAGGATCCGGCTGCCCTCCGCGAGCGGCTGGGGGGGGGGGGGGGGGG 120
QY 426 CTGGGGCTGCCCGGGAGACCGACAGAGAGCCGAGGAGCCCGCGGGAGGGGAGCTTT 485
Db 121 CTGGGGCTGCCCGGGAGACCGACAGAGAGCCCGAGGAGCCCGCGGGAGGGGAGCTTT 180
```


|||||
Db 1261 CATGTGCACCATGAGTTTCAGGACGGCATGGAAGGCCCTTTGTGCACCTTGGACATG 1320
Qy 1626 GAAGACTGTGGCTACACATTCACACACAGATCAGTCAACCCCTCATGACCATAGCCCTAT 1685
Db 1321 GAAGACTGTGGCTACACATTCACACACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380
Qy 1686 GTCATGTGCTGCCATTCGGCCCTCTTCATGTGCCACTCTGCCCTCATCTGGTGTGTCAGTGG 1745
Db 1381 GTCATGTGCTGCCATTCGGCCCTCTTCATGTGCCACTCTGCCCTCATCTGGTGTGTCAGTGG 1440
Qy 1746 CGTGTGCTCGCTGCTCGCCAGCAGCATGATGACATTTGCTGTGATGACATCTCCCTGCTG 1805
Db 1441 CGTGTGCTCGCTGCTCGCCAGCAGCATGATGACATTTGCTGTGATGACATCTCCCTGCTG 1500
Qy 1806 AAGTGAGGAGGCCCATCGGAGAGATAGACATTCCTCGTGGACACACACCTCCGCTGGTTCA 1865
Db 1501 AAGTGAGGAGGCCCATCGGAGAGATAGACATTCCTCGTGGACACACACCTCCGCTGGTTCA 1560
Qy 1866 CTTTGGTGCACAGTAGGAGACACAGATGGCACCTGTGTGCCAGAGCAGCTCAGGACCCCTCC 1925
Db 1561 CTTTGGTGCACAGTAGGAGACACAGATGGCACCTGTGTGCCAGAGCAGCTCAGGACCCCTCC 1620
Qy 1926 CCACCCACCAATGCCCTGTGCTTGTGAGGAGAGAAAGGCTGGCAAGGTGGGTTCAG 1985
Db 1621 CCACCCACCAATGCCCTGTGCTTGTGAGGAGAGAAAGGCTGGCAAGGTGGGTTCAG 1680
Qy 1986 GGACTGTACCTGTAGGAGACAGAAAGAGAAAGAGAGCACTCTGCTGGCGGGAATACT 2045
Db 1681 GGACTGTACCTGTAGGAGACAGAAAGAGAAAGAGCACTCTGCTGGCGGGAATACT 1740
Qy 2046 CTTGGTCACCTCAAAATTTAAGTCGGGAAATTCGCTGCTGAAACTTCAGCCCTGAACCT 2105
Db 1741 CTTGGTCACCTCAAAATTTAAGTCGGGAAATTCGCTGCTGAAACTTCAGCCCTGAACCT 1800
Qy 2106 TTGTCCACCAATTCCTTTAAATTCCTCAACCCAAAGTATTCCTTTCTTAAAGTTTCAGAA 2165
Db 1801 TTGTCCACCAATTCCTTTAAATTCCTCAACCCAAAGTATTCCTTTCTTAAAGTTTCAGAA 1860
Qy 2166 GTACTGGCATCACAGGAGTTACTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAG 2225
Db 1861 GTACTGGCATCACAGGAGTTACTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAG 1920
Qy 2226 AGACCAAGCTTTTCCCTGCTGCGCAAGTCAGTAGGAGAGATGACAGTTTGTCTATT 2285
Db 1921 AGACCAAGCTTTTCCCTGCTGCGCAAGTCAGTAGGAGAGATGACAGTTTGTCTATT 1980
Qy 2286 TGCTTTAGAGACAGGAGCTGTATAAACAAGCCTAACATTTGCTGCAAAAGATTGCCCTCTTGA 2345
Db 1981 TGCTTTAGAGACAGGAGCTGTATAAACAAGCCTAACATTTGCTGCAAAAGATTGCCCTCTTGA 2040
Qy 2346 ATT 2348
|||
Db 2041 ATT 2043

RESULT 8
AR178469
LOCUS AR178469 2541 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6319689.
ACCESSION AR178469
VERSION AR178469.1 GI:20219607
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2541)
AUTHORS Powell, D. J., Chapman, C. G., Murphy, K., and Smith, T. S.
TITLE US 6319689 A 1 20-NOV-2001;
JOURNAL Patent: US 6319689 A 1 20-NOV-2001;
FEATURES Location/Qualifiers
source 1. 2541
/organism="unknown"

BASE COUNT 598 a 673 c 675 g 579 t 16 others
ORIGIN
Query Match 86.2%; Score 2024; DB 6; Length 2541;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2038; Conservative 0; Mismatches 5; Indels 1; Gaps 1:
Qy 306 ATGGCCCAAGCCCTGCCCTGCTGCTGTGGATGGCGGGAGTGTCTGCTGCCAC 365
Db 1 ATGGCCCAAGCCCTGCCCTGCTGCTGTGGATGGCGGGAGTGTCTGCTGCCAC 60
Qy 366 GGCACCCAGCAGGATCCGGCTGCCCCTGCGCAGCGGCTGGGGGGGGCCCCCTGGGG 425
Db 61 GGCACCCAGCAGGATCCGGCTGCCCCTGCGCAGCGGCTGGGGGGGGCCCCCTGGGG 120
Qy 426 CTGGCGCTGCCCGGGAGACCGAGAGAGCCGAGGAGCCGCGCGGAGGCGAGCTTT 485
Db 121 CTGGCGCTGCCCGGGAGACCGAGAGAGCCGAGGAGCCGCGCGGAGGCGAGCTTT 180
Qy 486 GTGAGATGTGGACAACTTGAGGGGCAAGTCGGGGCAGGCTACTAGTGGAGATGACC 545
Db 181 GTGAGATGTGGACAACTTGAGGGGCAAGTCGGGGCAGGCTACTAGTGGAGATGACC 240
Qy 546 GTGGGACGCCCGCGCAGACGCTCAACATCTCTGTGTGGATACAGGACGAGTAACCTTGA 605
Db 241 GTGGGACGCCCGCGCAGACGCTCAACATCTCTGTGTGGATACAGGACGAGTAACCTTGA 300
Qy 606 GTGGGTGCTGCCCGCCCGCCCTCTCTGCTGCTACTACAGAGGACGCTGTGCCAGCACA 665
Db 301 GTGGGTGCTGCCCGCCCGCCCTCTCTGCTGCTACTACAGAGGACGCTGTGCCAGCACA 360
Qy 666 TACCGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAG 725
Db 361 TACCGGACCTCCGGAAGGGTGTGTATGAGCCCTACACCCAGGCAAGTGGGAAGGGAG 420
Qy 726 CTGGGACCGACCTGGTAAGCATCCCCCATGSCCCCAACGTCACGTGTGGTGGCCAACTT 785
Db 421 CTGGGACCGACCTGGTAAGCATCCCCCATGSCCCCAACGTCACGTGTGGTGGCCAACTT 480
Qy 786 GCTGCCATCACTGAATCAGACAAAGTTCTCATCAAGGCTCCAACTGGGAGGCACTCTG 845
Db 481 GCTGCCATCACTGAATCAGACAAAGTTCTCATCAAGGCTCCAACTGGGAGGCACTCTG 540
Qy 846 GGGCTGGCTTATGTGTAGATTGGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 905
Db 541 GGGCTGGCTTATGTGTAGATTGGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT 600
Qy 906 CTGTTAAAGCAGACCCAGTTCCTCCCAACCTCTTCTCCCTGCGAGCTTTGTGGTGGCTGCT 965
Db 601 CTGTTAAAGCAGACCCAGTTCCTCCCAACCTCTTCTCCCTGCGAGCTTTGTGGTGGCTGCT 660
Qy 966 CCCCTCAACCACTGTAAGTGTGGCTGCTGCGGAGGAGCATCATCATTTGGAGGTATC 1025
Db 661 CCCCTCAACCACTGTAAGTGTGGCTGCTGCGGAGGAGCATCATCATTTGGAGGTATC 720
Qy 1026 GACCACTGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGGTATTAT 1085
Db 721 GACCACTGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGGGGAGTGGTATTAT 780
Qy 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGACACTGCAAGGAG 1145
Db 781 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGACACTGCAAGGAG 840
Qy 1146 TACAACATGACAAGCATTTGTGGACAGTGGCAGCACCACCACTTCGTTTGGCCCAAGAA 1205
Db 841 TACAACATGACAAGCATTTGTGGACAGTGGCAGCACCACCACTTCGTTTGGCCCAAGAA 900
Qy 1206 GTGTTTGAAGCTGCAAGTCAAAATCCCATCAAGGAGGCTCTCCACGAGAGTTCCTGTAT 1265
Db 901 GTGTTTGAAGCTGCAAGTCAAAATCCCATCAAGGAGGCTCTCTCCACGAGAGTTCCTGTAT 960
Qy 1266 GGTGTTTGTGGTGTAGGAGAGCAGTGTGTGCTGGCAAGCAGGACCCACCTTTGGAACTT 1325
|||||

Db	1741	TCATTGGTCACCTCAAAATTTAAGTCGGGAAATTCCTGCTTGTAACACTTCAGCCCTCGAAC	1800
Qy	2105	TTTGTCACCACTTCCTTTAAATTCCTCAACCCAAAGATGATCTCTCTTTCTTAGTTTCAGA	2164
Db	1801	TTTGTCACCACTTCCTTTAAATTCCTCAACCCAAAGATGATCTCTCTTTCTTAGTTTCAGA	1860
Qy	2165	AGTACTGGCATCACGACGAGTTACCTTTGGGGTGTGTCCTGTGGTACCCCTGGGAGAGAA	2224
Db	1861	AGTACTGGCATCACGACGAGTTACCTTTGGGGTGTGTCCTGTGGTACCCCTGGGAGAGAA	1920
Qy	2225	GAGACCAAGCTGTGTTCCCTGCTGCGCAAAAGTCAAGTAGGAGAGGATGCACAGTTTGGCTAT	2284
Db	1921	GAGACCAAGCTGTGTTCCCTGCTGCGCAAAAGTCAAGTAGGAGAGGATGCACAGTTTGGCTAT	1980
Qy	2285	TTGCTTTAGACACAGGACTGTATATAACAAGCCCTAAACATTGCTCAAAAGATTGCTCTTTG	2344
Db	1981	TTGCTTTAGACACAGGACTGTATATAACAAGCCCTAAACATTGCTCAAAAGATTGCTCTTTG	2040
Qy	2345	AATT 2348	
Db	2041	AATT 2044	
RESULT 10			
LOCUS	E50816	Aspartate proteinase ASP2.	PAT 18-JUN-2001
DEFINITION	E50816	Aspartate proteinase ASP2	
ACCESSION	E50816	Aspartate proteinase ASP2	
VERSION	E50816.1	GI:13023199	
KEYWORDS	JP 2000060579-A/1.		
SOURCE	unidentified.		
ORGANISM	unidentified.		
REFERENCE	1. (bases 1 to 2541)		
AUTHORS	David,J.P., Conrad,G.C., Kay,M. and Trudy,S.S.		
TITLE	Aspartate proteinase ASP2		
JOURNAL	Patent JP 2000060579-A 1 29-FEB-2000;		
COMMENT	SMITKLINE-BEECHAM CORP		
OS	Unidentified		
PN	JP 2000060579-A/1		
PD	29-FEB-2000		
PF	03-AUG-1999 JP 1999219655		
PR	28-JAN-1997 GB 9701684.4		
PI	DAVID J POWER, CONRAD G CHAPPUMAN, KAY MAFI, TRUDY S SMITH		
PC	C12N15/09, A61K31/7088, A61K38/46, A61K39/00, A61K39/395, PC		
PC	A61K39/395, A61K48/00,		
PC	A61P25/28, A61P35/00, A61P43/00, C07K16/40, C12N1/15, C12N1/19, PC		
PC	C12N1/21,		
PC	C12N9/10, C12N9/64, C12Q1/37, G01N33/15, G01N33/50, G01N33/53, PC		
PC	G01N33/566,		
PC	G01N33/577//C12P21/08, C12N15/00, A61K37/54, C12N5/00 CC		
CC	Strandedness: Single;		
CC	Topology: Linear;		
FH	Key		
FT	source		
FT	Location/Qualifiers		
FT	1..2541		
FT	/organism='Unidentified'.		
FEATURES	source		
source	1..2541		
source	/organism='unidentified'		
source	/db_xref='taxon:32644'		
BASE COUNT	598 a 673 c 675 g 579 t 16 others		
ORIGIN			
Query Match	86.28;	Score 2024;	DB 6; Length 2541;
Best Local Similarity	99.76;	Pred. No. 0;	
Matches 2038:	Conservative	0; Mismatches	5; Indels 1; Gaps 1;
Qy	306	ATGGCCCAAGCCCTGCGCTCTCTGCTGTGGATGGCGGGGAGTGCTGCTGCCGCCAC	365
Db	1	ATGGCCCAAGCCCTGCGCTCTCTGCTGTGGATGGCGGGGAGTGCTGCTGCCGCCAC	60
Qy	366	GGCACCAGCAGCGCATCCGCTGCCCTGCGCGACGGCCCTGGGGGGCGCCCCCTGGGG	425

Db	61	GGCACCCAGCACGGCATCGGGTGGCCCTGCGCAGCGGCTGGGGGGCCCCCCCCCTGGGG	120
Qy	426	CTCGGGCTGCCCCGGGAGACCCAGAACGCCGAGAGCCCGGAGCGCCGGAGGGGACGCTTT	485
Db	121	CTCGGGCTGCCCCGGGAGACCCAGNAGACCCGAGAGCGCCGGAGGGGACGCTTT	180
Qy	486	GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGSCAGGGCTACTACGTGGAGATGACC	545
Db	181	GTGGAGATGGTGGACAACCTGAGGGCAAGTCGGGCGAGGCTACTACGTGGAGATGACC	240
Qy	546	GTGGGACGCCCCCGACGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA	605
Db	241	GTGGGACGCCCCCGACGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA	300
Qy	606	GTGGTGTCTGCCCCACCCCTTCCTGTGATCGCTACTACAGAGGCAGCTGTCCAGCAC	665
Db	301	GTGGTGTCTGCCCCACCCCTTCCTGTGATCGCTACTACAGAGGCAGCTGTCCAGCAC	360
Qy	666	TACCGGGACCTCCGGGAAGGGTGTGATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG	725
Db	361	TACCGGGACCTCCGGGAAGGGTGTGATGAGCCCTACACCCAGGGCAAGTGGGAAGGGGAG	420
Qy	726	CTGGGCACGGACCTGGTAAAGCATCCCCATGGCCCCAACGTCTACTGTGGTGCCACATT	785
Db	421	CTGGGCACGGACCTGGTAAAGCATCCCCATGGCCCCAACGTCTACTGTGGTGCCACATT	480
Qy	786	GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGAAGGCATCCCTG	845
Db	481	GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGAAGGCATCCCTG	540
Qy	846	GGGTGGGCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT	905
Db	541	GGGTGGGCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTGACTCT	600
Qy	906	CTGTTAAAGCAGACCCAGTTCCCAACCTTCTCTCCCTGCAGCTTTGTGGTGTGGCTTC	965
Db	601	CTGTTAAAGCAGACCCAGTTCCCAACCTTCTCTCCCTGCAGCTTTGTGGTGTGGCTTC	660
Qy	966	CCCCTCAACCACTGTAAGTGTGGCTCTGTCTGGAGGAGCATGATCATTTGAGGATATC	1025
Db	661	CCCCTCAACCACTGTAAGTGTGGCTCTGTCTGGAGGAGCATGATCATTTGAGGATATC	720
Qy	1026	GACCACTCGCTGTACACAGGCAGTCTCTGTTATACACCCATCCGGCGGGAGTGGTATTAT	1085
Db	721	GACCACTCGCTGTACACAGGCAGTCTCTGTTATACACCCATCCGGCGGGAGTGGTATTAT	780
Qy	1086	GAGTGTATCTTGC GG GTGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	1145
Db	781	GAGTGTATCTTGC GG GTGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	840
Qy	1146	TACAACTATGACAAGAGCATCTGACAGTGGCACACCACCACTTCGTTTGGCCCAAGAAA	1205
Db	841	TACAACTATGACAAGAGCATCTGACAGTGGCACACCACCACTTCGTTTGGCCCAAGAAA	900
Qy	1206	GTGTTTGAAGCTGCACTCAAAATCCATCAAGGCAGCCTCCTCCACGGAGAAAGTTCCCTGAT	1265
Db	901	GTGTTTGAAGCTGCACTCAAAATCCATCAAGGCAGCCTCCTCCACGGAGAAAGTTCCCTGAT	960
Qy	1266	GTTTCTCGCTAGGAGAGCAGCTGGTGTCTGGCAAGCAGGCACCACTTGGGAACATT	1325
Db	961	GTTTCTCGCTAGGAGAGCAGCTGGTGTCTGGCAAGCAGGCACCACTTGGGAACATT	1020
Qy	1326	TTCCAGCTCATCTCACTTACCTAATGGTGAGGTTACCAACCACTTCCTCCGCATCACC	1385
Db	1021	TTCCAGCTCATCTCACTTACCTAATGGTGAGGTTACCAACCACTTCCTCCGCATCACC	1080
Qy	1386	ATCCTTCGCGAGCAATACCTCGCGGCAGTGGAAAGATGTGGCCACGTCCTCCCAAGACGACTGT	1445
Db	1081	ATCCTTCGCGAGCAATACCTCGCGGCAGTGGAAAGATGTGGCCACGTCCTCCCAAGACGACTGT	1140
Qy	1446	TACAAGTTTGCCATCTCAGATCATCCACGGGACGTTTATGGGAGCTGTTATCATCTGGAG	1505
Db	1141	TACAAGTTTGCCATCTCAGATCATCCACGGGACGTTTATGGGAGCTGTTATCATCTGGAG	1200

Qy	1506	GGCTTCTACGTTGTCTTTGATCGGGGCCGAAACGAAATTTGGCTTTTGTGTGTCAGCGCTTGC	1566
Db	1201	GGCTTCTACGTTGTCTTTGATCGGGGCCGAAACGAAATTTGGCTTTTGTGTGTCAGCGCTTGC	1260
Qy	1566	CATGTGCACGATGAGTTTCAGGACGCGACGGGTGGAAGGCCCTTTTGTACCTTTGGACATG	1625
Db	1261	CATGTGCACGATGAGTTTCAGGACGCGACGGGTGGAAGGCCCTTTTGTACCTTTGGACATG	1320
Qy	1626	GAAGACTGTGGCTTACAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT	1685
Db	1321	GAAGACTGTGGCTTACAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT	1380
Qy	1686	GTTCATGGCTGCCATCTGCGGCCCTCTTTCATGCTGCCACTCTGCCCTCATGGTGTGTACGTGG	1745
Db	1381	GTTCATGGCTGCCATCTGCGGCCCTCTTTCATGCTGCCACTCTGCCCTCATGGTGTGTACGTGG	1440
Qy	1746	CGCTGCCCTCCGCTGCCCTGCGCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTGTG	1805
Db	1441	CGCTGCCCTCCGCTGCCCTGCGCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTGTG	1500
Qy	1806	AAGTGAAGGAGGCCATGGGCGAGAAGATAGAGATTCCGCTTGGAGCACACCTCCGTTGGTTTC	1864
Db	1501	AAGTGAAGGAGGCCATGGGCGAGAAGATAGAGATTCCGCTTGGAGCACACCTCCGTTGGTTTC	1560
Qy	1865	ACTTTGGTTCACAACTAGGAGACACAGATGGCACCTCTGTGGCCAGAGCACTTCAGGACCCCTC	1924
Db	1561	ACTTTGGTTCACAACTAGGAGACACAGATGGCACCTCTGTGGCCAGAGCACTTCAGGACCCCTC	1620
Qy	1925	CCCACCCACCAATTTGCCCTCTGCTTTGATGGAGAAAGGCTGGCAAGGTGGGTTCCTCA	1984
Db	1621	CCCACCCACCAATTTGCCCTCTGCTTTGATGGAGAAAGGCTGGCAAGGTGGGTTCCTCA	1680
Qy	1985	GGGACTGTACCTGTAGGAACAGAAAGAGAGAAAGCACTCTGCTGGCGGGAATAC	2044
Db	1681	GGGACTGTACCTGTAGGAACAGAAAGAGAGAAAGCACTCTGCTGGCGGGAATAC	1740
Qy	2045	TC TTGGTCACTCAAAATTTAAGTCGGGAATTTCTGCTGCTTGAAACTTCAGCCCTGAACC	2104
Db	1741	TC TTGGTCACTCAAAATTTAAGTCGGGAATTTCTGCTGCTTGAAACTTCAGCCCTGAACC	1800
Qy	2105	TTTTGTCCACCATTCCTTTAAATTTCTCAACCCCAAGATTTCTTTTTCTTTAGTTTCAGA	2164
Db	1801	TTTTGTCCACCATTCCTTTAAATTTCTCAACCCCAAGATTTCTTTTTCTTTAGTTTCAGA	1860
Qy	2165	AGTACTGGCATCACACGAGGTTTACCTTTGGCGTGTGTCCTGTGGTACCTTCGGCAGACAA	2224
Db	1861	AGTACTGGCATCACACGAGGTTTACCTTTGGCGTGTGTCCTGTGGTACCTTCGGCAGACAA	1920
Qy	2225	GAGACCAAGCTTGT TTTCCCTGCTGGCCCAAGTCAGTATGAGAGAGGATGACAGTTTGCCTAT	2284
Db	1921	GAGACCAAGCTTGT TTTCCCTGCTGGCCCAAGTCAGTATGAGAGAGGATGACAGTTTGCCTAT	1980
Qy	2285	TTGCTTTAGAGACAGGGACTGTATAAACAAAGCCCTAACATTTGGTGTCAAGATTTGCCTCTTG	2344
Db	1981	TTGCTTTAGAGACAGGGACTGTATAAACAAAGCCCTAACATTTGGTGTCAAGATTTGCCTCTTG	2040
Qy	2345	AATTT 2348	
Db	2041	AATTT 2044	

RESULT 11		
AX062111		
LOCUS	AX062111	3252 bp
DEFINITION	Sequence 1 from Patent WO0100655.	

REFERENCE 1 (bases 1 to 3252)
AUTHORS Tang, J. J., Hong, L., and Ghosh, A. K.
TITLE Inhibitors of cAMP-dependent protein kinase II and use thereof
PATENT WO 01/00665A (2001/01/04)
JOURNAL Oklahoma-Medical-Research-Foundation-(US) ; THE BOARD OF TRUSTEES
OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
source Location/Qualifiers
BASE COUNT 804 a 863 c 811 g 771 t 3 others
ORIGIN
Query Match 85.3% ; Score 2004 ; DB 6 ; Length 3252 ;
Best Local Similarity 100.0% ; Pred. No. 0 ;
Matches 2004 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;
QY 345 GCGGGAGTGTGCTGCCCGGACCCAGCAGCAGCGATCCGGCTGCCCTCGCAGCGGC 404
DB 1 GCGGGAGTGTGCTGCCCGGACCCAGCAGCAGCGATCCGGCTGCCCTCGCAGCGGC 60
QY 405 CTGGGGGCGCCCTCGGGCTCGGGCTGCCCGGAGACCGAGAGCCCGAGGAG 464
DB 61 CTGGGGGCGCCCTCGGGCTCGGGCTGCCCGGAGACCGAGAGCCCGAGGAG 120
QY 465 CCCGGCGGAGGGCAGCTTTGTGGAGATGCTGGACAACTGAGGGGCAAGTCGGGGCAG 524
DB 121 CCCGGCGGAGGGCAGCTTTGTGGAGATGCTGGACAACTGAGGGGCAAGTCGGGGCAG 180
QY 525 GGCTACTAGTGGAGATGACCGTGGGAGACCGCCCGGAGAGCGCTCAACATCCTGGTGGAT 584
DB 181 GGCTACTAGTGGAGATGACCGTGGGAGACCGCCCGGAGAGCGCTCAACATCCTGGTGGAT 240
QY 585 ACAGCAGCAGTAATTTGCAAGTGGTGTGCGCCCGGACCCCTTCCCTGCATCGCTACTAC 644
DB 241 ACAGCAGCAGTAATTTGCAAGTGGTGTGCGCCCGGACCCCTTCCCTGCATCGCTACTAC 300
QY 645 CAGAGGAGCTGTCCAGCACATACCGGAGCTCCGGAGGCTGTATGTGCGCCCTACACC 704
DB 301 CAGAGGAGCTGTCCAGCACATACCGGAGCTCCGGAGGCTGTATGTGCGCCCTACACC 360
QY 705 CAGGGCAAGTGGGAAGGGAGCTGGGACACCGACCTGTGAACATCCCGGAGGCGCCCAAC 764
DB 361 CAGGGCAAGTGGGAAGGGAGCTGGGACACCGACCTGTGAACATCCCGGAGGCGCCCAAC 420
QY 765 GTCACTGTGCGTGCACATATGCTGCCATCACTGAATCAGACAAAGTCTTTCATCAACGGC 824
DB 421 GTCACTGTGCGTGCACATATGCTGCCATCACTGAATCAGACAAAGTCTTTCATCAACGGC 480
QY 825 TCCAACTGGGAAGGATCCTCGGGCTGGCCCTATGCTGAGATTGCCAGGCTGAGACATCC 884
DB 481 TCCAACTGGGAAGGATCCTCGGGCTGGCCCTATGCTGAGATTGCCAGGCTGAGACATCC 540
QY 885 CTGAGGCTTTCTTTGACTCTCTGTTAAAGCAGACCCACGCTTCCCAACCTCTTCCCTG 944
DB 541 CTGAGGCTTTCTTTGACTCTCTGTTAAAGCAGACCCACGCTTCCCAACCTCTTCCCTG 600
QY 945 CAGCTTTGTGCTGGCTTCCCTCCCAACCAAGTCTGAAGTGTGCGCTCTGTGGAGGG 1004
DB 601 CAGCTTTGTGCTGGCTTCCCTCCCAACCAAGTCTGAAGTGTGCGCTCTGTGGAGGG 660
QY 1005 AGCATGATCATTTGGAGGTATCGACCACTCGCTGTACAGGAGCTCTCTGGTATACACC 1064
DB 661 AGCATGATCATTTGGAGGTATCGACCACTCGCTGTACAGGAGCTCTCTGGTATACACC 720
QY 1065 ATCCGGGGGAGTGGTATTATGAGTGTATGTTGCGGGGTGGAGATCAATGGACAGAT 1124
DB 721 ATCCGGGGGAGTGGTATTATGAGTGTATGTTGCGGGGTGGAGATCAATGGACAGAT 780
QY 1125 CTGAAATGGAGTCAAGGAGTACAACCTATCACAAGAGCATTTGGAGAGTGGGACCAACC 1184
DB 781 CTGAAATGGAGTCAAGGAGTACAACCTATCACAAGAGCATTTGGAGAGTGGGACCAACC 840

QY 1185 AACCTTCGTTTCCCAAGAAAGTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCCCTCC 1244
DB 841 AACCTTCGTTTCCCAAGAAAGTGTTTGAAGCTGCAGTCAATCCATCAAGCAGCCCTCC 900
QY 1245 TCCACGGAGAGTTCCTCGTATGTTTCTGGCTAGGAGACAGCTGGTGTCTGGCAAGCA 1304
DB 901 TCCACGGAGAGTTCCTCGTATGTTTCTGGCTAGGAGACAGCTGGTGTCTGGCAAGCA 960
QY 1305 GGCACACCCCTTGGAAACATTTTCCAGTCTATCTCAGTCTACCTAATGGGTGAGTTACC 1364
DB 961 GGCACACCCCTTGGAAACATTTTCCAGTCTATCTCAGTCTACCTAATGGGTGAGTTACC 1020
QY 1365 AACAGTCTTCCGATCACCATCTTCCGAGCAATACCTGCGGCCAGTGGAAAGTGG 1424
DB 1021 AACAGTCTTCCGATCACCATCTTCCGAGCAATACCTGCGGCCAGTGGAAAGTGG 1080
QY 1425 GCCAGTCCCAAGAGAGCTGTTACAAAGTTCGCAATCTCAGTCTACCTGCGGCCAGTGGT 1484
DB 1081 GCCAGTCCCAAGAGAGCTGTTACAAAGTTCGCAATCTCAGTCTACCTGCGGCCAGTGGT 1140
QY 1485 ATGGGAGCTGTTATCATGGAGGCTTCTACGTTTGTCTTGTATCGGGCCCGAAACGAATT 1544
DB 1141 ATGGGAGCTGTTATCATGGAGGCTTCTACGTTTGTCTTGTATCGGGCCCGAAACGAATT 1200
QY 1545 GGCTTTGCTGTACGGCTTGCATGTGCACGATGATGATTCAGGACGCGAGCGTGGAAAGC 1604
DB 1201 GGCTTTGCTGTACGGCTTGCATGTGCACGATGATGATTCAGGACGCGAGCGTGGAAAGC 1260
QY 1605 CTTTGTGTCACCTTGGACATGGAAGACTGTGGCTACACATTCACACAGATGAGTCA 1664
DB 1261 CTTTGTGTCACCTTGGACATGGAAGACTGTGGCTACACATTCACACAGATGAGTCA 1320
QY 1665 ACCCTCATGACCATAGCCTATGTCATGCTGCATCTCGCCCTCTTCTATGCTGCCACTC 1724
DB 1321 ACCCTCATGACCATAGCCTATGTCATGCTGCATCTCGCCCTCTTCTATGCTGCCACTC 1380
QY 1725 TGCCTCATGCTGTGCTAGTGGCGCTGCTCGCTGCTCGCCAGCAGCATGATGACTTT 1784
DB 1381 TGCCTCATGCTGTGCTAGTGGCGCTGCTCGCTGCTCGCCAGCAGCATGATGACTTT 1440
QY 1785 GCTGATGACATCTCCCTGCTGAGTGAAGAGAGGCCCATGGGAGAGATGAGATTCCTCT 1844
DB 1441 GCTGATGACATCTCCCTGCTGAGTGAAGAGAGGCCCATGGGAGAGATGAGATTCCTCT 1500
QY 1845 GGACACACCTCCGCTGGTTCATCTTGGTCACAAGTAGAGACACAGATGGCAGCTGTGGC 1904
DB 1501 GGACACACCTCCGCTGGTTCATCTTGGTCACAAGTAGAGACACAGATGGCAGCTGTGGC 1560
QY 1905 CAGAGCAGCTCAGGACCTCCCGACCCACCAATGCTCTGCTTGTGAGGAGGAGAAA 1964
DB 1561 CAGAGCAGCTCAGGACCTCCCGACCCACCAATGCTCTGCTTGTGAGGAGGAGAAA 1620
QY 1965 GGCTGGCAAGTGGTTCACAGGAGCTGACCTGTAGGAACACAGAAAGAGAAAGAG 2024
DB 1621 GGCTGGCAAGTGGTTCACAGGAGCTGACCTGTAGGAACACAGAAAGAGAAAGAG 1680
QY 2025 CACTCTGCTGGCGGGAATACCTTGGTCACTCAAAATTAAGTCGGGAAATTCCTGCTGT 2084
DB 1681 CACTCTGCTGGCGGGAATACCTTGGTCACTCAAAATTAAGTCGGGAAATTCCTGCTGT 1740
QY 2085 TGAACATTCAGCCCTGAAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT 2144
DB 1741 TGAACATTCAGCCCTGAAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT 1800
QY 2145 CTTCTTTTCTTAGTTTTCAGAGTACTGGCATCACAGCAGGTTACTTTGGCGTGTGCTCC 2204
DB 1801 CTTCTTTTCTTAGTTTTCAGAGTACTGGCATCACAGCAGGTTACTTTGGCGTGTGCTCC 1860
QY 2205 TGTGTACCTCGGCAAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA 2264
DB 1861 TGTGTACCTCGGCAAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA 1920
QY 2265 GAGGATGCACAGTTTGTCTTTAGCTTTAGAGACAGGGGACTGTATAAACAGCCTAACATT 2324

Db 1921 GAGGATGCACAGTTTGTCTATTGCTTTAGAGACAGGGACTGTATAAAGCCTAACATT 1980
QY 2325 GGTGCAAGATTGCGCTTCTTGAATT 2348
Db 1981 GGTGCAAGATTGCGCTTCTTGAATT 2004

RESULT 12
AX063201 LOCUS AX063201 3252 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 1 from Patent WO0100663.
ACCESSION AX063201
VERSION AX063201.1 GI:12541045
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3252)
AUTHORS Tang, J. J., Lin, X. and Koelsch, G.
TITLE Catalytically active recombinant memapsin and methods of use there
of
JOURNAL Patent: WO 0100663-A 1 04-JAN-2001;
Oklahoma Medical Research Foundation (US)
FEATURES
source 1..3252
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 804 a 863 c 811 g 771 t 3 others
ORIGIN

Query Match 85.3%; Score 2004; DB 6; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 CGGGAGTGTCTGCCGCCACGGCACCAGCAGGCGATCCGGCTGCCCTCGCGACGGCG 404
Db 1 CGGGAGTGTCTGCCGCCACGGCACCAGCAGGCGATCCGGCTGCCCTCGCGACGGCG 60
QY 405 CTGG 464
Db 61 CTGG 120
QY 465 CCGGGCCGGAGGGGAGCTTTGTGGAGATGTGGACAACTGTAGGGGCAAGTCGGGGCAG 524
Db 121 CCGGGCCGGAGGGGAGCTTTGTGGAGATGTGGACAACTGTAGGGGCAAGTCGGGGCAG 180
QY 525 GGCTACTACGTGGAGATGACCGTGGGAGCCCGCCCGCAGACGCTCAACATCCTGTGGAT 584
Db 181 GGCTACTACGTGGAGATGACCGTGGGAGCCCGCCCGCAGACGCTCAACATCCTGTGGAT 240
QY 585 ACAGGCACGATTAATTTGCAGTGGGTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 644
Db 241 ACAGGCACGATTAATTTGCAGTGGGTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 300
QY 645 CAGAGGAGCTGTCCAGCACATACCGGGACCTCGGGAAGGGTGTGTATGCGCCCTACACC 704
Db 301 CAGAGGAGCTGTCCAGCACATACCGGGACCTCGGGAAGGGTGTGTATGCGCCCTACACC 360
QY 705 CAGGGCAAGTGGGAAGGGAGTGGGACCGACCTGGTAAAGCATCCCGCCATGGCCCCCAAC 764
Db 361 CAGGGCAAGTGGGAAGGGAGTGGGACCGACCTGGTAAAGCATCCCGCCATGGCCCCCAAC 420
QY 765 GTCACTGTGGTGGCAACATTTGCTGCATCAGTGAATCAGACAGTTTCTTCATCAACGGC 824
Db 421 GTCACTGTGGTGGCAACATTTGCTGCATCAGTGAATCAGACAAAGTTCTTCATCAACGGC 480
QY 825 TCCAACTGGGAAGGGATCCTGGGGCTGGCTATGCTCAGATTGCCAGGCTGACGACTCC 884
Db 481 TCCAACTGGGAAGGGATCCTGGGGCTGGCTATGCTCAGATTGCCAGGCTGACGACTCC 540
QY 885 CTGGAGGCTTTCTTGACTCTCTGGTAAAGACAGACCCACGTTCCCAACCTCTCTCCCTG 944

Db 541 CTGGAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCTG 600
QY 945 CAGCTTTTGGTGCTGGCTTCCCTTCAACAGTCTGAAGTCTGGCTCTCTGTCGGAGG 1004
Db 601 CAGCTTTTGGTGCTGGCTTCCCTTCAACAGTCTGAAGTCTGGCTCTCTGTCGGAGG 660
QY 1005 AGCATGATCATTCGAGGTATCGACCATCTGCTGTACACAGGAGTCTCTGCTATACACC 1064
Db 661 AGCATGATCATTCGAGGTATCGACCATCTGCTGTACACAGGAGTCTCTGCTATACACC 720
QY 1065 ATCCGCGGGAGTGGTATTTAGGTGATCATTTGTCGGGGTGGAGATCAATGGACAGAT 1124
Db 721 ATCCGCGGGAGTGGTATTTAGGTGATCATTTGTCGGGGTGGAGATCAATGGACAGAT 780
QY 1125 CTGAAATGGACTGCAAGAGGTACAATATGACAGAGCATTTGTGACAGTGGCACACC 1184
Db 781 CTGAAATGGACTGCAAGAGGTACAATATGACAGAGCATTTGTGACAGTGGCACACC 840
QY 1185 AACCTTCTGTTTCCCAAGAAAGTGTGGAAGCTGCAGTCAAAATCCATCAAGGAGCTCC 1244
Db 841 AACCTTCTGTTTCCCAAGAAAGTGTGGAAGCTGCAGTCAAAATCCATCAAGGAGCTCC 900
QY 1245 TCCACGGAGAAGTTCCTGTATGTTTCTGGCTAGGAGAGCAGTGTGTCTGGCAAGCA 1304
Db 901 TCCACGGAGAAGTTCCTGTATGTTTCTGGCTAGGAGAGCAGTGTGTCTGGCAAGCA 960
QY 1305 GGCACACCCCTTGGAAACATTTTCCAGTCACTCTCACATTAATGGTGAAGTTTACC 1364
Db 961 GGCACACCCCTTGGAAACATTTTCCAGTCACTCTCACATTAATGGTGAAGTTTACC 1020
QY 1365 AACCACTCTTCCGATCACCATCTTCCGAGCAATACCTCGGGCAGTGGGAAGATGTG 1424
Db 1021 AACCACTCTTCCGATCACCATCTTCCGAGCAATACCTCGGGCAGTGGGAAGATGTG 1080
QY 1425 GCCACCTCCCAAGACGACTGTTACAAGTTTGCACATCTCACAGTCACTCCAGGGCATT 1484
Db 1081 GCCACCTCCCAAGACGACTGTTACAAGTTTGCACATCTCACAGTCACTCCAGGGCATT 1140
QY 1485 ATGGAGCTGTTTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCGAAACGATT 1544
Db 1141 ATGGAGCTGTTTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCGAAACGATT 1200
QY 1545 GCTTTGCTGTACAGCGCTTGCATGTGCACATGATGATTCAGGACGCGGTGGAAGGC 1604
Db 1201 GCTTTGCTGTACAGCGCTTGCATGTGCACATGATGATTCAGGACGCGGTGGAAGGC 1260
QY 1605 CCTTTTGCACCTTGGACATGGAAGACTGTGGCTACAACATTTCCACAGACAGATGACTCA 1664
Db 1261 CCTTTTGCACCTTGGACATGGAAGACTGTGGCTACAACATTTCCACAGACAGATGACTCA 1320
QY 1665 ACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTCGGCCCTCTTTCATGCTGCCACTC 1724
Db 1321 ACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTCGGCCCTCTTTCATGCTGCCACTC 1380
QY 1725 TCCTTCATGCTGTGTCAGTGGCGCTGCCCTGCCCTGCCAGCAGCATGATGACTTTT 1784
Db 1381 TCCTTCATGCTGTGTCAGTGGCGCTGCCCTGCCCTGCCAGCAGCATGATGACTTTT 1440
QY 1785 GCTGATGACATCTCCCTTGTGAGTGGAGGGCCCATGGGCAAGATAGAGATTTCCCTT 1844
Db 1441 GCTGATGACATCTCCCTTGTGAGTGGAGGGCCCATGGGCAAGATAGAGATTTCCCTT 1500
QY 1845 GGACACACCTCGGTGTTCACTTTGGTCACAAGTAGAGACACAGATGGCACCTGTGGC 1904
Db 1501 GGACACACCTCGGTGTTCACTTTGGTCACAAGTAGAGACACAGATGGCACCTGTGGC 1560
QY 1905 CAGAGCACCTCAGAGCCCTCCCAACCCCAAAATCCCTTCTGCTTCTGATGGAGAGGAAA 1964
Db 1561 CAGAGCACCTCAGAGCCCTCCCAACCCCAAAATCCCTTCTGCTTCTGATGGAGAGGAAA 1620
QY 1965 GCTGSCAAGGTGGTTCAGGGGACTGTACCTGTAGGAACAGAAAGAGAGAAAGAAAG 2024

Db 1621 GCGTGCAGAGTGGGTTCACGAGGACTGTACCTGTAGGAACAGAAAAGAGAAAGAAG 1680
QY 2025 CACTCTGCTGCGGGGAATACTCTTGTGTCACTCACTCAAAATTTAAGTCGGGAAATTCGTCTGCT 2084
Db 1681 CACTCTGCTGCGGGGAATACTCTTGTGTCACTCAAAATTTAAGTCGGGAAATTCGTCTGCT 1740
QY 2085 TGAACCTTCAGCCCTGAAACCTTTGTGTCACCAATTCCTTTAAATTCCTCAACCCAAAGTATT 2144
Db 1741 TGAACCTTCAGCCCTGAAACCTTTGTGTCACCAATTCCTTTAAATTCCTCAACCCAAAGTATT 1800
QY 2145 CTTCTTTTCTTATGTTTCAGAACTACTGGCATCACAGCAGGTTACCTTGGCGTGTCTCC 2204
Db 1801 CTTCTTTTCTTATGTTTCAGAACTACTGGCATCACAGCAGGTTACCTTGGCGTGTCTCC 1860
QY 2205 TGTGTACCTTGGCAGAGAGAACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA 2264
Db 1861 TGTGTACCTTGGCAGAGAGAACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA 1920
QY 2265 GAGGATGCACAGTTTGTCTATTGCTTTTAGAGACAGGAGCTGTATAAACAAAGCCTAACATT 2324
Db 1921 GAGGATGCACAGTTTGTCTATTGCTTTTAGAGACAGGAGCTGTATAAACAAAGCCTAACATT 1980
QY 2325 GGTGCAAGATTGCCCTCTTGAATT 2348
Db 1981 GGTGCAAGATTGCCCTCTTGAATT 2004

RESULT 13
AX472368
LOCUS AX472368 3252 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 1 from Patent WO02053594.
ACCESSION AX472368
VERSION AX472368.1 GI:22207364
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ghosh,A.K., Koelsch,G. and Tang,J.J.
TITLE Inhibitors of menapsin-2 and use thereof
JOURNAL Patent: WO/02053594-A 1 11-JUL-2002;
OKLAHOMA MED RES FOUND (US); TRUSTEES OF THE UNIVERSITY OF ILLINOIS
(US)
FEATURES
Location/Qualifiers
1..3252
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 804 a 863 c 811 g 771 t 3 others
ORIGIN
Query Match 85.3%; Score 2004; DB 6; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 GCGGAGTGTGCTGCCACGCGACCCAGCAGCGCATCCGGCTGCCCTGCGCAGCGGC 404
Db 1 GCGGAGTGTGCTGCCACGCGACCCAGCAGCGCATCCGGCTGCCCTGCGCAGCGGC 60
QY 405 CTGGGGGGCCCCCTGGGGCTGCGGGCTGCCCGGGAGACCGAGAGGCCGAGGAG 464
Db 61 CTGGGGGGCCCCCTGGGGCTGCGGGCTGCCCGGGAGACCGAGAGGCCGAGGAG 120
QY 465 CCCGGCGGAGGGGAGCTTTGTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAG 524
Db 121 CCCGGCGGAGGGGAGCTTTGTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAG 180
QY 525 GGCTACTAGCTGGAGATGACCGTGGCGAGCCCCCGCAGACGCTCAACATCTCTGGTGGAT 584
Db 181 GGCTACTAGCTGGAGATGACCGTGGCGAGCCCCCGCAGACGCTCAACATCTCTGGTGGAT 240
QY 585 ACAGGACGAGTAACCTTTGCAGTGGGTGCTGCCCCCAGCCCTTCTGTCATCCTACTAC 644
Db 1321 ACCCTCATGACCATAGCCTATGTATGGCTGCCATCTGCGCCCTCTTCTATGCTGCCACTC 1380

Db 241 ACAGGACGAGTAACATTTGCAGTGGGTGTGCCCCCAGCCCTTCTCTGTCATCGCTACTAC 300
QY 645 CAGAGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGCTGTGTATGTGCGCTACACC 704
Db 301 CAGAGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGCTGTGTATGTGCGCTACACC 360
QY 705 CAGGCAAGTGGGAAGGGAGCTGGGCACCGACCTTGGTAAGCATGCCCTATGCCCCCAAC 764
Db 361 CAGGCAAGTGGGAAGGGAGCTGGGCACCGACCTTGGTAAGCATGCCCTATGCCCCCAAC 420
QY 765 GTCACTGTGGTGGTCCCAACATTTGCTGCATCACTGAATCAGACAAGTTCTTCATCAACGGC 824
Db 421 GTCACTGTGGTGGTCCCAACATTTGCTGCATCACTGAATCAGACAAGTTCTTCATCAACGGC 480
QY 825 TCCAACTGGGAAGGCACTTCTGGGCTTCCCTTCAACCACTGTGAGATTGCCAGGCTGACGACTCC 884
Db 481 TCCAACTGGGAAGGCACTTCTGGGCTTCCCTTCAACCACTGTGAGATTGCCAGGCTGACGACTCC 540
QY 885 CTGGAGCCTTTCTTTGACTCTCTGTTAAAGCAGACCCACCTTCCCAACCTTCTTCTCCCTG 944
Db 541 CTGGAGCCTTTCTTTGACTCTCTGTTAAAGCAGACCCACCTTCCCAACCTTCTTCTCCCTG 600
QY 945 CAGCTTTGGTGGTCTGGCTTCCCTTCAACCACTGTGAGATTGCCAGGCTGACGAGGG 1004
Db 601 CAGCTTTGGTGGTCTGGCTTCCCTTCAACCACTGTGAGATTGCCAGGCTGACGAGGG 660
QY 1005 AGCATGATCATTCGAGGTATCGACCACTCGCTGTACACAGGCACTCTCTGTTATACACC 1064
Db 661 AGCATGATCATTCGAGGTATCGACCACTCGCTGTACACAGGCACTCTCTGTTATACACC 720
QY 1065 ATCCGGCGGAGTGGTATTATGAGTGTATTTGCGGGTGGAGATCAATGACAGGAT 1124
Db 721 ATCCGGCGGAGTGGTATTATGAGTGTATTTGCGGGTGGAGATCAATGACAGGAT 780
QY 1125 CTGAAAATGGAGTGCAGGAGTACACTATGACAAGAGCATTTGGACAGTGGCACACC 1184
Db 781 CTGAAAATGGAGTGCAGGAGTACACTATGACAAGAGCATTTGGACAGTGGCACACC 840
QY 1185 AACCTTCTGTTGCCCAAGAAAGTGTGGAAGTGCAGTCAAAATCCATCAAGGAGCCTCC 1244
Db 841 AACCTTCTGTTGCCCAAGAAAGTGTGGAAGTGCAGTCAAAATCCATCAAGGAGCCTCC 900
QY 1245 TCCACGAGAGTTCCTCTGATGTTTCTGCTAGGAGAGCAGCTGGTGTCTGGCAAGCA 1304
Db 901 TCCACGAGAGTTCCTCTGATGTTTCTGCTAGGAGAGCAGCTGGTGTCTGGCAAGCA 960
QY 1305 GGCACACCCCTTGGAAACATTTTCCAGTCACTCTACCTTAATGGGTGAGTTTACC 1364
Db 961 GGCACACCCCTTGGAAACATTTTCCAGTCACTCTACCTTAATGGGTGAGTTTACC 1020
QY 1365 AACCACTTCTCCGATCACCATCTTCCGACCAATACCTGCGGCGAGTGGAGATGTG 1424
Db 1021 AACCACTTCTCCGATCACCATCTTCCGACCAATACCTGCGGCGAGTGGAGATGTG 1080
QY 1425 GCCACGTCCCAAGACGACTGTTTAAAGTTCCTTACAGTCTTACAGTCACTCCAGGGCAGCTTT 1484
Db 1081 GCCACGTCCCAAGACGACTGTTTAAAGTTCCTTACAGTCTTACAGTCACTCCAGGGCAGCTTT 1140
QY 1485 ATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGATGGGCCCCGAAACGAATT 1544
Db 1141 ATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGATGGGCCCCGAAACGAATT 1200
QY 1545 GGCCTTGTCTCAGCGCTTCCCATGTCCAGATGCTTACAGGAGCGGCGGAGGAGGC 1604
Db 1201 GGCCTTGTCTCAGCGCTTCCCATGTCCAGATGCTTACAGGAGCGGCGGAGGAGGC 1260
QY 1605 CCTTTTGTCACTTGGACATGGAAGCTGTGGCTTACAACTTCCACAGACAGATGAGTCA 1664
Db 1261 CCTTTTGTCACTTGGACATGGAAGCTGTGGCTTACAACTTCCACAGACAGATGAGTCA 1320
QY 1665 ACCCTCATGACCATAGCCTATGTATGGCTGCCATCTGCGCCCTCTTCTATGCTGCCACTC 1724
Db 1321 ACCCTCATGACCATAGCCTATGTATGGCTGCCATCTGCGCCCTCTTCTATGCTGCCACTC 1380

QY	1725	TGCTCATGTGTGTCAGTGGCGTGCCTCGCTCGCTCGCCAGCAGCATGATGACTTT	1784	/db_xref="taxon:9606"
Db	1381	TGCTCATGTGTGTCAGTGGCGTGCCTCGCTCGCCAGCAGCATGATGACTTT	1440	<1..11467
QY	1785	GCTGATGACATCTCCCTGCTGTAAGTGAAGAGGCCATGGCAGAGATAGATTTCCCTT	1844	/function="cleaves beta-secretase site of beta-amyloid precursor protein"
Db	1441	GCTGATGACATCTCCCTGCTGTAAGTGAAGAGGCCATGGCAGAGATAGATTTCCCTT	1500	/note="membrane-associated aspartic protease 2"
QY	1845	GGACACACCTCGTGTCTCACTTTGCTCACAAGTAGGAGACACAGATGGCCTGTGC	1904	/codon_start=1
Db	1501	GGACACACCTCGTGTCTCACTTTGCTCACAAGTAGGAGACACAGATGGCCTGTGC	1560	/product="memapsin 2"
QY	1905	CAGAGCACCTCAGACACCTCCCCACCCACCAATGCCCTTGCTGATGGAGAAGAAAA	1964	/protein_id="AAPI3715.1"
Db	1561	CAGAGCACCTCAGACACCTCCCCACCCACCAATGCCCTTGCTGATGGAGAAGAAAA	1620	/db_xref="GI:6470293"
QY	1965	GGCTGCAAGGTGGTTCAGGAGTGTACCTGTAGGAACAAGAAAGAGAAGAAG	2024	/translation="AGVLPAGTQHGIRLPRLRSLGGLGAPLRLRLPHRETDDEEPGR
Db	1621	GGCTGCAAGGTGGTTCAGGAGTGTACCTGTAGGAACAAGAAAGAGAAGAAG	1680	GSFVEMVNDLRKSGQGYVEMTGPSPOTLNLVDTSNSFANCAAPFLHRYTOR
QY	2025	CACCTGCTGGCGGAATACCTTTGGTCACCTCAAAATTAAGTCGGGAAATTCGTGCT	2084	QUSSTYDLRKGVYVPTQKWEGLDLSIHPGNVTVRANTAAIINQSKFFTING
Db	1681	CACCTGCTGGCGGAATACCTTTGGTCACCTCAAAATTAAGTCGGGAAATTCGTGCT	1740	SNWEGILGLAYAEIARPDLSLEPFSLVKQTHVNLFSQLQCGAGFLINQSLVAS
QY	2085	TGAACCTTACGCCCTGAACCTTTGCTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT	2144	GSMIIIGTDHSLYTGSLWTPIRREWYEVITVRVEINGDLKMDCKEYNDKSIDV
Db	1741	TGAACCTTACGCCCTGAACCTTTGCTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT	1800	SGTTLNLRPKKVFEEAAVKSTKAASSTKFPDGLVEOLVCHQAGTTPNNIPVISLY
QY	2145	CTTCTTTTCTTAGTTTCAGAACTACTGGCATCACAGCAGGTTACCTTGGCGTGTGCC	2204	LMGEVNTQSFRTILPOOYLRPVEDVATSDDDCYKFAJSQSTGTVMGAVIMEGFYV
Db	1801	CTTCTTTTCTTAGTTTCAGAACTACTGGCATCACAGCAGGTTACCTTGGCGTGTGCC	1860	FRARAKRIFAVSACHVHDEFTAAVEGPFVTLDMEDCGTNIPTQDDESTLMTIAYVNA
QY	2205	TGTGGTACCTTGGCAGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA	2264	AICALFMLPLCLMVQWRCRLRQQHDFADFDISLLK"
Db	1861	TGTGGTACCTTGGCAGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA	1920	BASE COUNT 804 a 863 c 812 g 773 t
QY	2265	GAGGATGCACAGTTGCTATTGCTTTAGACAGAGGACTGTATAACAAAGCCTAACATT	2324	ORIGIN
Db	1921	GAGGATGCACAGTTGCTATTGCTTTAGACAGAGGACTGTATAACAAAGCCTAACATT	1980	Query Match 85.3%; Score 2004; DB 9; Length 3252;
QY	2325	GGTGCAGAGATGGCTCTTGAATT 2348		Best Local Similarity 100.0%; Pred. No. 0;
Db	1981	GGTGCAGAGATGGCTCTTGAATT 2004		Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
LOCUS	AF200193	3252 bp mRNA, linear	PRI 16-FEB-2000	
DEFINITION	Homo sapiens memapsin 2 mRNA, partial cds.			
ACCESSION	AF200193			
VERSION	AF200193.1	GI:6470292		
KEYWORDS	Homo sapiens.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
AUTHORS	1 (bases 1 to 3252)			
TITLE	Lin.X., Koelsch,G., Wu,S., Downs,D., Dashti,A. and Tang,J.			
JOURNAL	Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein			
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1456-1460 (2000)			
PUBMED	20144060			
REFERENCE	2 (bases 1 to 3252)			
AUTHORS	Lin.X., Koelsch,G. and Tang,J.			
TITLE	Direct Submission			
JOURNAL	Submitted (28-Oct-1999) Protein Studies Program, Oklahoma Medical Research Foundation, 825 N.E. 13th Street, Oklahoma City, OK 73104, USA			
FEATURES	Location/Qualifiers			
Source	1..3252			
	/organism="Homo sapiens"			

Db	1801	CTTCTTTTCTTAGTTTTCAGAAAGTACTGGCATCACACGAGGTACTCTTGGCGTGTGTCCC	1860
QY	2205	TGTGGTACCCCTGGCAGAGAGAGACCAAGCTTGTTCCTCTGCTGGCCCAAGTCAAGTAGGA	2264
Db	1861	TGTGGTACCCCTGGCAGAGAGAGACCAAGCTTGTTCCTCTGCTGGCCCAAGTCAAGTAGGA	1920
QY	2265	GAGGATGCACAGTTTCTATTGCTTTAGAGACAGGGACTGTATATAAACAAGCCTAAACATT	2324
Db	1921	GAGGATGCACAGTTTCTATTGCTTTAGAGACAGGGACTGTATATAAACAAGCCTAAACATT	1980
QY	2325	GGTGCAAGATTGCCCTCTTGAATT	2348
Db	1981	GGTGCAAGATTGCCCTCTTGAATT	2004
RESULT 15			
AX105387			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1966; Conservative			
QY	306	ATGGCCCAAGCCTGCCCCTGGCTCTGCTGTGATGGCGGGAGTCTGCTGCCCTGCCAC	365
Db	1	ATGGCCCAAGCCTGCCCCTGGCTCTGCTGTGATGGCGGGAGTCTGCTGCCCTGCCAC	60
QY	366	GGCACCACGACGCGCATCCGGCTGCCCCTGCGCAGCGGCTGGGGGGCGCCCTTGGGG	425
Db	61	GGCACCACGACGCGCATCCGGCTGCCCCTGCGCAGCGGCTGGGGGGCGCCCTTGGGG	120
QY	426	CTCGGGCTGCCCGGGAGACCGACGAAGCCCGAGGAGCCCGGGCGGGGAGCGCTTT	485
Db	121	CTCGGGCTGCCCGGGAGACCGACGAAGCCCGAGGAGCCCGGGCGGGGAGCGCTTT	180
QY	486	GTGGAGATGTGGACAACTTGAGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC	545
Db	181	GTGGAGATGTGGACAACTTGAGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC	240
QY	546	GTGGGAGCCCCCGGCAGACGCTCAACATCTCTGGTGGATACAGGCAGCAGTAACCTTGTGA	605
Db	241	GTGGGAGCCCCCGGCAGACGCTCAACATCTCTGGTGGATACAGGCAGCAGTAACCTTGTGA	300
QY	606	GTGGGTGCTGCCCGCCACCCCTTCTGCAATCGCTACTACAGAGGACGCTGTCCAGCACA	665
Db	301	GTGGGTGCTGCCCGCCACCCCTTCTGCAATCGCTACTACAGAGGACGCTGTCCAGCACA	360
QY	666	TACCGGACCTCCGGGAAGGTTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG	725
Db	361	TACCGGACCTCCGGGAAGGTTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG	420
QY	726	CTGGGACCGACCTGGTAAAGCATTCGCCCATGCCCCCAACGCTCACTGTGCGTGGCAACATT	785

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:00:56 ; Search time 17 Seconds
(without alignments)
789.227 Million cell updates/sec

Title: US-09-723-722A-43

Perfect score: 2419
Sequence: 1 ETDEPEPGRGSGFVEMVD.....CLRLRQHQHDFADDISLIK 456

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued_Patents_AA.*
- 1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
 - 2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
 - 3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
 - 4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
 - 5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*
 - 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2419	100.0	501	US-09-548-372D-4	Sequence 4, Appli
2	2419	100.0	501	US-09-548-367D-4	Sequence 4, Appli
3	2413	99.8	501	US-09-009-191-2	Sequence 2, Appli
4	2390	98.8	501	US-09-713-158-2	Sequence 2, Appli
5	2390	98.8	501	US-09-548-372D-8	Sequence 8, Appli
6	2390	98.8	501	US-09-548-367D-8	Sequence 8, Appli
7	2320	95.9	774	US-09-009-191-4	Sequence 4, Appli
8	2264.5	93.6	476	US-09-548-372D-6	Sequence 6, Appli
9	2264.5	93.6	476	US-09-548-367D-6	Sequence 6, Appli
10	2238.5	92.5	476	US-09-548-372D-73	Sequence 73, Appli
11	2238.5	92.5	476	US-09-548-367D-73	Sequence 73, Appli
12	2160	89.3	453	US-09-548-372D-30	Sequence 30, Appli
13	2160	89.3	453	US-09-548-367D-30	Sequence 30, Appli
14	2160	89.3	459	US-09-548-372D-32	Sequence 32, Appli
15	2160	89.3	459	US-09-548-367D-32	Sequence 32, Appli
16	2155	89.1	433	US-09-548-372D-26	Sequence 26, Appli
17	2155	89.1	433	US-09-548-367D-26	Sequence 26, Appli
18	2135	89.1	446	US-09-548-372D-22	Sequence 22, Appli
19	2135	89.1	446	US-09-548-367D-22	Sequence 22, Appli
20	2155	89.1	459	US-09-548-372D-24	Sequence 24, Appli
21	2155	89.1	459	US-09-548-367D-24	Sequence 24, Appli
22	2090.5	86.4	425	US-09-548-372D-28	Sequence 28, Appli
23	2090.5	86.4	425	US-09-548-367D-28	Sequence 28, Appli
24	2005.5	82.9	428	US-09-548-372D-51	Sequence 51, Appli
25	2005.5	82.9	428	US-09-548-367D-51	Sequence 51, Appli
26	2005.5	82.9	434	US-09-548-372D-53	Sequence 53, Appli
27	2005.5	82.9	434	US-09-548-367D-53	Sequence 53, Appli

28	1171.5	48.4	518	3	US-08-999-723-2	Sequence 2, Appli
29	1171.5	48.4	518	4	US-09-434-427-2	Sequence 2, Appli
30	1171.5	48.4	518	4	US-09-548-372D-2	Sequence 2, Appli
31	1171.5	48.4	518	4	US-09-548-367D-2	Sequence 2, Appli
32	1146.5	47.4	514	4	US-09-717-432-2	Sequence 2, Appli
33	1146.5	47.4	514	4	US-09-912-484-2	Sequence 2, Appli
34	299.5	12.4	396	1	US-08-208-007A-13	Sequence 13, Appli
35	299.5	12.4	396	4	US-09-032-523-9	Sequence 9, Appli
36	299.5	12.4	396	4	US-08-915-095A-13	Sequence 13, Appli
37	299.5	12.4	396	4	US-08-798-096-13	Sequence 13, Appli
38	299.5	12.4	396	4	US-08-798-095A-13	Sequence 13, Appli
39	295.5	12.2	412	1	US-08-208-007A-12	Sequence 12, Appli
40	295.5	12.2	412	4	US-08-974-691-4	Sequence 4, Appli
41	295.5	12.2	412	4	US-08-915-095A-12	Sequence 12, Appli
42	295.5	12.2	412	4	US-08-798-096-12	Sequence 12, Appli
43	295.5	12.2	412	4	US-08-798-095A-12	Sequence 12, Appli
44	279.5	11.6	458	6	5217891-15	Patent No. 5217891
45	273.5	11.3	409	1	US-09-640-305-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-548-372D-4
; Sequence 4, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-4

Query Match	100.0%;	Score 2419;	DB 4;	Length 501;
Best Local Similarity	100.0%;	Pred. No. 1.1e-251;		
Matches 456;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ETDEPEPGRGSGFVEMVDNLKRGSGQYVEMTVGSPPTNLILVDTGSSNFAVGAAP	60	
Db	46	ETDEPEPGRGSGFVEMVDNLKRGSGQYVEMTVGSPPTNLILVDTGSSNFAVGAAP	105	
Qy	61	HPFLHRYQRLSSTYRDLRKGVVYPTQGWEGELGTLVSIHPGPNVTVRANIAAITE	120	
Db	106	HPFLHRYQRLSSTYRDLRKGVVYPTQGWEGELGTLVSIHPGPNVTVRANIAAITE	165	
Qy	121	SDKPFFINGSNWEGLGLAYAEIARPDSDSLEPFFDSLVRKQTHVHPNLSLQLCGAGFPLNQS	180	
Db	166	SDKPFFINGSNWEGLGLAYAEIARPDSDSLEPFFDSLVRKQTHVHPNLSLQLCGAGFPLNQS	225	
Qy	181	EVLASVGSMIIIGIDHSLSLTGSLWYTPIRREWYVEYIIVRVEINGODLKMCKEYNYDK	240	
Db	226	EVLASVGSMIIIGIDHSLSLTGSLWYTPIRREWYVEYIIVRVEINGODLKMCKEYNYDK	285	
Qy	241	SIVDSGTTNLRLPKKVFEAAVKSIAASSTKEKFPDGFGLGVLGVCWQAGTTPWNIIPVIS	300	
Db	286	SIVDSGTTNLRLPKKVFEAAVKSIAASSTKEKFPDGFGLGVLGVCWQAGTTPWNIIPVIS	345	

QY 301 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYV 360
DB 346 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYV 405
QY 361 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 420
DB 406 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 465
QY 421 CALFMLPLCLMWQWRCRLRQHQHDDFADDISLLK 456
DB 466 CALFMLPLCLMWQWRCRLRQHQHDDFADDISLLK 501

RESULT 2

US-09-548-367D-4
; Sequence 4, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-4

Query Match 100.0%; Score 2419; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.le-251;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEEPPGRGGSFVEMVDNLRGKSGGYVYVMTVGPQTNLILVDTGSSNFAVGAAP 60
DB 46 ETDEEPEEPPGRGGSFVEMVDNLRGKSGGYVYVMTVGPQTNLILVDTGSSNFAVGAAP 105
QY 61 HPFLHRYYQRLSSTYRDLRKGVVYPTTQGWEGELGTDLSVPHGPNVTVRANIAAITE 120
DB 106 HPFLHRYYQRLSSTYRDLRKGVVYPTTQGWEGELGTDLSVPHGPNVTVRANIAAITE 165
QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDLVKQTHVPHNLSLQCGAGFPNQS 180
DB 166 SDRFFINGSNWEGILGLAYAEIARPDSDLVKQTHVPHNLSLQCGAGFPNQS 225
QY 181 EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 240
DB 226 EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIAASSTEFKPDGFWLGEQLVCWQAGTTPWNIFPVIS 300
DB 286 SIVDSGTTNLRPKKVFEEAAVKSIAASSTEFKPDGFWLGEQLVCWQAGTTPWNIFPVIS 345
QY 301 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYV 360
DB 346 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYV 405
QY 361 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 420
DB 406 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 465

QY 421 CALFMLPLCLMWQWRCRLRQHQHDDFADDISLLK 456
DB 466 CALFMLPLCLMWQWRCRLRQHQHDDFADDISLLK 501

RESULT 3

US-09-009-191-2
; Sequence 2, Application US/09009191
; Patent No. 6315689
; GENERAL INFORMATION:
; APPLICANT: POWELL, DAVID
; APPLICANT: CHAPMAN, CONRAD
; APPLICANT: MURPHY, KAY
; APPLICANT: SMITH, TRUDI
; TITLE OF INVENTION: ASP2
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,191
; FILING DATE: 20-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 9701684.4
; FILING-DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70368
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 845169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-009-191-2

Query Match 99.8%; Score 2413; DB 4; Length 501;
Best Local Similarity 99.8%; Pred. No. 4.8e-251;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEEPEEPPGRGGSFVEMVDNLRGKSGGYVYVMTVGPQTNLILVDTGSSNFAVGAAP 60
DB 46 ETDEEPEEPPGRGGSFVEMVDNLRGKSGGYVYVMTVGPQTNLILVDTGSSNFAVGAAP 105
QY 61 HPFLHRYYQRLSSTYRDLRKGVVYPTTQGWEGELGTDLSVPHGPNVTVRANIAAITE 120
DB 106 HPFLHRYYQRLSSTYRDLRKGVVYPTTQGWEGELGTDLSVPHGPNVTVRANIAAITE 165
QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDLVKQTHVPHNLSLQCGAGFPNQS 180
DB 166 SDRFFINGSNWEGILGLAYAEIARPDSDLVKQTHVPHNLSLQCGAGFPNQS 225
QY 181 EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 240
DB 226 EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIAASSTEFKPDGFWLGEQLVCWQAGTTPWNIFPVIS 300

;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;; FILE OF INVENTION: THEREOF
;; FILE REFERENCE: 29915/6280H
;; CURRENT APPLICATION NUMBER: US/09/548,367D
;; CURRENT FILING DATE: 2000-04-12
;; PRIOR APPLICATION NUMBER: US 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: US 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: US 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 8
;; LENGTH: 501
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-548-367D-8

Query Match 98.8%; Score 2390; DB 4; Length 501;
Best Local Similarity 98.2%; Pred. No. 1.5e-248;
Matches 448; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGRGSGFYEMVDNLKSGQGYVEMTVGSPQTLNLLVDTGSSNFVGAAP 60
DB 46 ETDEEPEEPRGRGSGFYEMVDNLKSGQGYVEMTVGSPQTLNLLVDTGSSNFVGAAP 105
QY 61 HPLHRYQRLSSTYRDLKGVVYPTGKWEGLGTLVSPHGPNTVVRANIAAITE 120
DB 106 HPLHRYQRLSSTYRDLKGVVYPTGKWEGLGTLVSPHGPNTVVRANIAAITE 165
QY 121 SOKFFINGSNWEGILGAYAEIARPDSSLEPFDSLVKQTHVPHNLSLQCGAGFPLNQS 180
DB 166 SOKFFINGSNWEGILGAYAEIARPDSSLEPFDSLVKQTHVPHNLSLQCGAGFPLNQT 225
QY 181 EVLASVGGSMITGGIDHSLYTGSLWTPTRREWYEVIIIVRVEINGQDLKMDCKEYNYDK 240
DB 226 EALASVGGSMITGGIDHSLYTGSLWTPTRREWYEVIIIVRVEINGQDLKMDCKEYNYDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVSKAASSTKFPDGFGLGQVLCVQAGTTPWNIFFVLS 300
DB 286 SIVDSGTTNLRPKKVFEEAAVSKAASSTKFPDGFGLGQVLCVQAGTTPWNIFFVLS 345
QY 301 LYLMGEVTTNQSFRTITLPOQYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYV 360
DB 346 LYLMGEVTTNQSFRTITLPOQYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYV 405
QY 361 FDRARRKRGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DB 406 FDRARRKRGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMVQWRCRLCRHQHDDFADDISLLK 456
DB 466 CALFMLPLCLMVQWRCRLCRHQHDDFADDISLLK 501

RESULT 7
US-09-009-191-4
; Sequence 4, Application US/09009191
; Patent No. 6319689
; GENERAL INFORMATION:
; APPLICANT: POWELL, DAVID
; APPLICANT: CHAPMAN, CONRAD
; APPLICANT: MURPHY, KAY
; APPLICANT: SMITH, TRUDI
; TITLE OF INVENTION: ASP2
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE

;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19482
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/009,191
;; FILING DATE: 20-JAN-1998
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: UK 9701684.4
;; FILING DATE: 28-JAN-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: PRESTIA, PAUL F
;; REGISTRATION NUMBER: 23,031
;; REFERENCE/DOCKET NUMBER: GH-70368
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-407-0700
;; TELEFAX: 610-407-0701
;; TELEX: 846169
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 774 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-009-191-4

Query Match 95.9%; Score 2320; DB 4; Length 774;
Best Local Similarity 98.9%; Pred. No. 1e-240;
Matches 439; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 GSFVEMVDNLKSGQGYVEMTVGSPQTLNLLVDTGSSNFVGAAPPHFLHRYQRL 72
DB 1 GSFVEMVDNLKSGQGYVEMTVGSPQTLNLLVDTGSSNFVGAAPPHFLHRYQRL 60
QY 73 SSTYRDLKGVVYPTGKWEGLGTLVSPHGPNTVVRANIAAITEADKFFINGSNME 132
DB 61 SSTYRDLKGVVYPTGKWEGLGTLVSPHGPNTVVRANIAAITEADKFFINGSNME 120
QY 133 GILGLAYAEIARPDSSLEPFDSLVKQTHVPHNLSLQCGAGFPLNQSSEVLASVGGSMII 192
DB 121 GILGLAYAEIARPDSSLEPFDSLVKQTHVPHNLSLQCGAGFPLNQSSEVLASVGGSMII 180
QY 193 GGTIDHSLYTGSLWTPTRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLR 252
DB 181 GGTIDHSLYTGSLWTPTRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLR 240
QY 253 PKKVFEEAAVSKAASSTKFPDGFGLGQVLCVQAGTTPWNIFFVLSLYLMGEVTTNQSF 312
DB 241 PKKVFEEAAVSKAASSTKFPDGFGLGQVLCVQAGTTPWNIFFVLSLYLMGEVTTNQSF 300
QY 313 RITILPQOYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYVDFRARRKRGFAV 372
DB 301 RITILPQOYLRPVEDVATSDQDCYKFAISQSSGTGTVMGAVIMEGFYVDFRARRKRGFAV 360
QY 373 SACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMV 432
DB 361 SACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMV 420
QY 433 COMRCLRLCRHQHDDFADDISLLK 456
DB 421 COMRCLRLCRHQHDDFADDISLLK 444

RESULT 8
US-09-548-372D-6
; Sequence 6, Application US/09548372D
; Patent No. 6420534

```

|||||
286 SIVDSGTTNLRPKKVFEEAAVKSIRAASTKFPDGFNLGELVLCWQAGTTPNNIFPVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
|||||
346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 405
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
|||||
406 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMVQWRCRLRCLRQHQHDDFADDISLLK 456
|||||
466 CALFMLPLCLMVQWRCRLRCLRQHQHDDFADDISLLK 501

RESULT 4
US-09-713-158-2
; Sequence 2, Application US/09713158
; Patent No. 6361975
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: LI, XIAOTONG
; APPLICANT: POWELL, DAVID J.
; APPLICANT: CHRISTIE, GARY
; TITLE OF INVENTION: MOUSE ASPARTIC SECRETASE-2 (MASP-2)
; FILE REFERENCE: GP-70660
; CURRENT APPLICATION NUMBER: US/09/713,158
; PRIOR FILING DATE: 2000-11-15
; PRIOR FILING DATE: 2000-11-15
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-09-713-158-2

Query Match 98.8%; Score 2390; DB 4; Length 501;
Best Local Similarity 98.2%; Pred. No. 1.5e-248;
Matches 448; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 ETDEEPEPGRGSRGSEFVMDNLRGSGQGYVYVEMTVGSPPTNLNLDVDTGSSNFVAVGAAP 60
|||||
46 ETDEEPEPGRGSRGSEFVMDNLRGSGQGYVYVEMTVGSPPTNLNLDVDTGSSNFVAVGAAP 105
QY 61 HPFLHRYQRLSSTYRDLRKGVYVPTQGWEGELGTLVSIPIHGPNTVVRANIAATE 120
|||||
106 HPFLHRYQRLSSTYRDLRKGVYVPTQGWEGELGTLVSIPIHGPNTVVRANIAATE 165
QY 121 SDKEFFINGSNWEGILGLAYAEIARPDSDLEPFDSLVKQTHVFNPLFSLQCCGAGFPLNQS 180
|||||
166 SDKEFFINGSNWEGILGLAYAEIARPDSDLEPFDSLVKQTHVFNPLFSLQCCGAGFPLNQS 225
QY 181 EVLASVGGSMIIGGDHSLYTGSLWYTPIRREWYVEVIVRVEINGQDLKMDCKEYNYDK 240
|||||
226 EALASVGGSMIIGGDHSLYTGSLWYTPIRREWYVEVIVRVEINGQDLKMDCKEYNYDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIRAASTKFPDGFNLGELVLCWQAGTTPNNIFPVIS 300
|||||
286 SIVDSGTTNLRPKKVFEEAAVKSIRAASTKFPDGFNLGELVLCWQAGTTPNNIFPVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
|||||
346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 405
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
|||||
406 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMVQWRCRLRCLRQHQHDDFADDISLLK 456
|||||
466 CALFMLPLCLMVQWRCRLRCLRQHQHDDFADDISLLK 501

```

```

Db 466 CALFMLPLCLMVQWRCRLRCLRQHQHDDFADDISLLK 501

RESULT 5
US-09-548-372D-8
; Sequence 8, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-548-372D-8

Query Match 98.8%; Score 2390; DB 4; Length 501;
Best Local Similarity 98.2%; Pred. No. 1.5e-248;
Matches 448; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 ETDEEPEPGRGSRGSEFVMDNLRGSGQGYVYVEMTVGSPPTNLNLDVDTGSSNFVAVGAAP 60
|||||
46 ETDEEPEPGRGSRGSEFVMDNLRGSGQGYVYVEMTVGSPPTNLNLDVDTGSSNFVAVGAAP 105
QY 61 HPFLHRYQRLSSTYRDLRKGVYVPTQGWEGELGTLVSIPIHGPNTVVRANIAATE 120
|||||
106 HPFLHRYQRLSSTYRDLRKGVYVPTQGWEGELGTLVSIPIHGPNTVVRANIAATE 165
QY 121 SDKEFFINGSNWEGILGLAYAEIARPDSDLEPFDSLVKQTHVFNPLFSLQCCGAGFPLNQS 180
|||||
166 SDKEFFINGSNWEGILGLAYAEIARPDSDLEPFDSLVKQTHVFNPLFSLQCCGAGFPLNQS 225
QY 181 EVLASVGGSMIIGGDHSLYTGSLWYTPIRREWYVEVIVRVEINGQDLKMDCKEYNYDK 240
|||||
226 EALASVGGSMIIGGDHSLYTGSLWYTPIRREWYVEVIVRVEINGQDLKMDCKEYNYDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIRAASTKFPDGFNLGELVLCWQAGTTPNNIFPVIS 300
|||||
286 SIVDSGTTNLRPKKVFEEAAVKSIRAASTKFPDGFNLGELVLCWQAGTTPNNIFPVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
|||||
346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 405
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
|||||
406 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMVQWRCRLRCLRQHQHDDFADDISLLK 456
|||||
466 CALFMLPLCLMVQWRCRLRCLRQHQHDDFADDISLLK 501

RESULT 6
US-09-548-367D-8
; Sequence 8, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.

```

2/10/1999


```

; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-6

Query Match      93.6%; Score 2264.5; DB 4; Length 476;
Best Local Similarity 94.5%; Pred. No. 4.4e-235;
Matches 431; Conservative 0; Mismatches 0; Indels 25; Gaps 1;

QY 1 ETDEEPEEPCRRGSGFVEMVDNLGRKSGQGYVYVEMTVGSPPTLNILVDTGSSNFVAGAAP 60
Db 46 ETDEEPEEPCRRGSGFVEMVDNLGRKSGQGYVYVEMTVGSPPTLNILVDTGSSNFVAGAAP 105

QY 61 HPFLHRYQRQLSSTYRDRLRGKYVYPYTOGKWEGLGTLVSIHPGPNVTVRANIAAITE 120
Db 106 HPFLHRYQRQLSSTYRDRLRGKYVYPYTOGKWEGLGTLVSIHPGPNVTVRANIAAITE 165

QY 121 SKRFFINGSNWEGILGLAYAEIARPDPSLEPPFDSLKVQTHVPLNLSQLCGAGFPLNQS 180
Db 166 SKRFFINGSNWEGILGLAYAEIAR-----LCGAGFPLNQS 200

QY 181 EVLASVGGSMIIGGDHSLYTGSLWYTPIRREWYEVYIIVRVEINGODLKMCKEYNYDK 240
Db 201 EVLASVGGSMIIGGDHSLYTGSLWYTPIRREWYEVYIIVRVEINGODLKMCKEYNYDK 260

QY 241 SIVDSTTNLRLPKKVFEEAAVSIKAASTKEKPPDGFNLGEQLVVCWQAGTTPWNIFFVIS 300
Db 261 SIVDSTTNLRLPKKVFEEAAVSIKAASTKEKPPDGFNLGEQLVVCWQAGTTPWNIFFVIS 320

QY 301 LYLMEVNTNOSFRITILPOOYLRPVEDVATSDDDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 321 LYLMEVNTNOSFRITILPOOYLRPVEDVATSDDDCYKFAISQSSTGTVMGAVIMEGFYV 380

QY 361 FDRARRKRGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
Db 381 FDRARRKRGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 440

QY 421 CALFMPLPLCLMVQWRCRLCRLROHQHDDFADDISLLK 456
Db 441 CALFMPLPLCLMVQWRCRLCRLROHQHDDFADDISLLK 476

RESULT 9
US-09-548-372D-6
; Sequence 6, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23

```

```

; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-6

Query Match      93.6%; Score 2264.5; DB 4; Length 476;
Best Local Similarity 94.5%; Pred. No. 4.4e-235;
Matches 431; Conservative 0; Mismatches 0; Indels 25; Gaps 1;

QY 1 ETDEEPEEPCRRGSGFVEMVDNLGRKSGQGYVYVEMTVGSPPTLNILVDTGSSNFVAGAAP 60
Db 46 ETDEEPEEPCRRGSGFVEMVDNLGRKSGQGYVYVEMTVGSPPTLNILVDTGSSNFVAGAAP 105

QY 61 HPFLHRYQRQLSSTYRDRLRGKYVYPYTOGKWEGLGTLVSIHPGPNVTVRANIAAITE 120
Db 106 HPFLHRYQRQLSSTYRDRLRGKYVYPYTOGKWEGLGTLVSIHPGPNVTVRANIAAITE 165

QY 121 SKRFFINGSNWEGILGLAYAEIARPDPSLEPPFDSLKVQTHVPLNLSQLCGAGFPLNQS 180
Db 166 SKRFFINGSNWEGILGLAYAEIAR-----LCGAGFPLNQS 200

QY 181 EVLASVGGSMIIGGDHSLYTGSLWYTPIRREWYEVYIIVRVEINGODLKMCKEYNYDK 240
Db 201 EVLASVGGSMIIGGDHSLYTGSLWYTPIRREWYEVYIIVRVEINGODLKMCKEYNYDK 260

QY 241 SIVDSTTNLRLPKKVFEEAAVSIKAASTKEKPPDGFNLGEQLVVCWQAGTTPWNIFFVIS 300
Db 261 SIVDSTTNLRLPKKVFEEAAVSIKAASTKEKPPDGFNLGEQLVVCWQAGTTPWNIFFVIS 320

QY 301 LYLMEVNTNOSFRITILPOOYLRPVEDVATSDDDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 321 LYLMEVNTNOSFRITILPOOYLRPVEDVATSDDDCYKFAISQSSTGTVMGAVIMEGFYV 380

QY 361 FDRARRKRGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
Db 381 FDRARRKRGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 440

QY 421 CALFMPLPLCLMVQWRCRLCRLROHQHDDFADDISLLK 456
Db 441 CALFMPLPLCLMVQWRCRLCRLROHQHDDFADDISLLK 476

RESULT 10
US-09-548-372D-73
; Sequence 73, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73

```


Db 166 SDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 225
QY 181 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 240
Db 226 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 285
QY 241 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 300
Db 286 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
Db 346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 405
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 408
Db 406 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 453
RESULT 13
US-09-548-367D-30
; Sequence 30, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-30
Query Match 89.3%; Score 2160; DB 4; Length 453;
Best Local Similarity 100.0%; Pred. No. 7.4e-224;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPPGRRGSGFVEMVDNLRKSGQGYVEMTVGSPPTNLILVDTGSSNFAVGAAP 60
Db 46 ETDEEPEEPPGRRGSGFVEMVDNLRKSGQGYVEMTVGSPPTNLILVDTGSSNFAVGAAP 105
QY 61 HPFLHRYRQRLSSTYRDLRKGVVYPYTOGKWEGLGTDLVSIHPGPNVTVRANIAAITE 120
Db 106 HPFLHRYRQRLSSTYRDLRKGVVYPYTOGKWEGLGTDLVSIHPGPNVTVRANIAAITE 165
QY 121 SKFFFTNGSNWEGILGLAYAEIARPDSDLEPPFDSLQKTHVNLFSLQLCGAGFPLNQS 180
Db 166 SDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 225
QY 181 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 240
Db 226 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 285
QY 241 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 300
Db 286 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
Db 346 SDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 408
QY 181 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 240
Db 226 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 285
QY 241 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 300
Db 286 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360

Db 346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 405
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 408
Db 406 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 453
RESULT 14
US-09-548-372D-32
; Sequence 32, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-32
Query Match 89.3%; Score 2160; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 7.5e-224;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPPGRRGSGFVEMVDNLRKSGQGYVEMTVGSPPTNLILVDTGSSNFAVGAAP 60
Db 46 ETDEEPEEPPGRRGSGFVEMVDNLRKSGQGYVEMTVGSPPTNLILVDTGSSNFAVGAAP 105
QY 61 HPFLHRYRQRLSSTYRDLRKGVVYPYTOGKWEGLGTDLVSIHPGPNVTVRANIAAITE 120
Db 106 HPFLHRYRQRLSSTYRDLRKGVVYPYTOGKWEGLGTDLVSIHPGPNVTVRANIAAITE 165
QY 121 SDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 180
Db 166 SDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 225
QY 181 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 240
Db 226 EVLASVGGSMITGGIDHSLYTGSLWTPPIRREWYEVIIVRVEINGODLKMDCKEYNDK 285
QY 241 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 300
Db 286 SIVDSGTTNLRPKPKVFEAAVKSIIKAASSTKPPDGFWLGEQLVCWQAGTTPWNIPFVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
Db 346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 405
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 408
Db 406 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTD 453
RESULT 15
US-09-548-367D-32
; Sequence 32, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-32

Query Match 89.3%; Score 2160; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 7.5e-224;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ETDEEPEEPCRRGSFVEMVDNLRKSGQYVVTGSPQTLNLLVDTGSSNFAVGAAP	60
Db	46	ETDEEPEEPCRRGSFVEMVDNLRKSGQYVVTGSPQTLNLLVDTGSSNFAVGAAP	105
Qy	61	HPFLHRYRQRLSSTYRDLRKGVVYPTQKWEGLGTLVSIHPGPNVTVRANIAATE	120
Db	106	HPFLHRYRQRLSSTYRDLRKGVVYPTQKWEGLGTLVSIHPGPNVTVRANIAATE	165
Qy	121	SKFFINGSNWEGILGLAYAEIARPDLSLEPFDLSLVKQTHVPNLFSLQLCGAGPPLNOS	180
Db	166	SKFFINGSNWEGILGLAYAEIARPDLSLEPFDLSLVKQTHVPNLFSLQLCGAGPPLNOS	225
Qy	181	EVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDK	240
Db	226	EVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDK	285
Qy	241	SIVDSGTTNLRPKKVFEEAAVSKIAASSTKFPDGFMLGEQLVCWQAGTTPWNIFFPVIS	300
Db	286	SIVDSGTTNLRPKKVFEEAAVSKIAASSTKFPDGFMLGEQLVCWQAGTTPWNIFFPVIS	345
Qy	301	LYLMGEVTHQSFRITILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVV	360
Db	346	LYLMGEVTHQSFRITILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVV	405
Qy	361	FDRARRKRGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDES	408
Db	406	FDRARRKRGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDES	453

Search completed: March 4, 2003, 10:01:52
Job time : 19 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2003, 13:37:20 ; Search time 324.5 seconds
(without alignments)
16294.891 Million cell updates/sec

Title: US-09-723-722A-44
Perfect score: 2348
Sequence: 1 ccacgcggccctcacagc.....caagaattgcctcttgatt 2348

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2347.6	100.0	2348	21	AAA59551 DNA encoding a hum
2	2336.4	99.5	16080	21	AAA59553 DNA clone pCEK Cl.
3	2181.2	92.9	5757	24	ABL39774 Human NS cDNA sequ
4	2039.8	86.9	2070	21	AAAL5662 Human aspartyl pro
5	2039.8	86.9	2070	22	AAD17865 Human aspartyl pro
6	2039.8	86.9	2070	22	AAD13021 Human aspartyl pro
7	2039.8	86.9	2070	22	AAAL5662 Human cDNA encodin
8	2039.8	86.9	2070	22	AAAL5662 Human cDNA encodin
9	2039.8	86.9	2070	22	AAD06739 Human aspartyl pro

10	2039.8	86.9	2070	24	ABL52457 Human Asp-2(a) nuc
11	2039.8	86.9	2070	24	ABL49914 Human aspartyl pro
12	2024	86.2	2541	19	AAV41696 Nucleotide sequenc
13	2004	85.3	3252	22	AAF31848 Human memapsin 2 c
14	2004	85.3	3252	22	AAF28101 Memapsin 2 DNA.. H
15	2004	85.3	3252	24	ABK88641 cDNA encoding huma
16	1881	80.1	1977	21	AAAL5663 Human aspartyl pro
17	1881	80.1	1977	22	AAD17866 Human aspartyl pro
18	1881	80.1	1977	22	AAD13022 Human cDNA encodin
19	1881	80.1	1977	22	AAAL5662 Human cDNA encodin
20	1881	80.1	1977	22	AAAL5662 Human cDNA encodin
21	1881	80.1	1977	22	AAAL5662 Human cDNA encodin
22	1881	80.1	1977	22	AAAL5662 Human cDNA encodin
23	1881	80.1	1977	24	ABL49915 Human aspartyl pr
24	1835.4	78.2	2370	19	AAV41697 Partial nucleotide
25	1792	76.3	1979	22	AAK94824 Human full-length
26	1648	70.2	1747	20	AAK97602 Extended human sec
27	1550	66.0	2043	21	AAAL5664 Murine aspartyl pr
28	1550	66.0	2043	22	AAD17867 Murine aspartyl pr
29	1550	66.0	2043	22	AAD13023 Murine aspartyl pr
30	1550	66.0	2043	22	AAAL5662 Mouse cDNA encodin
31	1550	66.0	2043	22	AAAL5662 DNA encoding mouse
32	1550	66.0	2043	22	AAAL5662 Murine aspartyl pr
33	1511.2	64.4	2158	24	ABK63758 Mouse Asp-2(a) nuc
34	1511.2	64.4	2158	24	ABK63758 Rat sequence diffe
35	1504.4	64.1	2907	23	AAK82237 Human encoding novel
36	1503	64.0	1503	21	AAK28278 Human cDNA encodin
37	1503	64.0	1503	21	AAK59550 DNA encoding a hum
38	1503	64.0	1527	24	ABA02406 FLAG-tagged human
39	1476.2	62.9	1911	22	AAK09485 Human aspartyl pro
40	1423	60.6	2914	23	AAK73798 DNA encoding novel
41	1355.8	57.7	1362	21	AAAL5668 Modified human asp
42	1355.8	57.7	1362	22	AAD17878 Human-Asp 2(a) pro
43	1355.8	57.7	1362	22	AAD13034 Human-Asp2(a) delt
44	1355.8	57.7	1362	22	AAAL5662 Human cDNA encodin
45	1355.8	57.7	1362	22	AAAL5662 DNA encoding human

ALIGNMENTS

RESULT 1
AAA59551
ID AAA59551 standard; DNA; 2348 BP.

XX	AC	AAA59551;			
XX	DT	14-NOV-2000 (first entry)			
XX	DE	DNA encoding a human beta-secretase enzyme.			
XX	DE	Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;			
XX	KW	amyloid plaque component; Alzheimer's disease; amyloidogenic disease;			
XX	KW	inhibitor; ss.			
XX	OS	Homo sapiens.			
XX	PH	Key	Location/Qualifiers		
XX	FT	CDS	306..1811		
XX	FT		/*tag= a		
XX	FT		/product= "beta-secretase"		
XX	PN	WO200047618-A2.			
XX	PD	17-AUG-2000.			
XX	PF	10-FEB-2000; 2000WO-US03819.			
XX	PR	10-FEB-1999; 99US-0119571.			
XX	PR	15-JUN-1999; 99US-0139172.			
XX	PA	(ELAN-) ELAN PHARM INC.			

PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX WPI: 2000-533011/48.
DR P-PSDB; AAB07896.
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
XX Disclosure; Fig 1B; 121pp; English.
XX
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence encodes a human beta-secretase enzyme.
XX
SQ Sequence 2348 BP; 489 A; 713 C; 661 G; 484 T; 1 other;

Query Match 100.0%; Score 2347.6; DB 21; Length 2348;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCGCGCCCTCACAGCCCGCGGAGCCCGAGCCGCGTGCACAGCTGCGCGCGG 60
DB 1 CCATGCGCGCCCTCACAGCCCGCGGAGCCCGAGCCGCGTGCACAGCTGCGCGCGG 60

QY 61 CSFTGCCGATGAGCGGCTCGGATCCAGCCTCTCCCTGCTCCCGTGTCTGCGGAT 120
DB 61 CSFTGCCGATGAGCGGCTCGGATCCAGCCTCTCCCTGCTCCCGTGTCTGCGGAT 120

QY 121 CTCCCTGTACCGCTCTCCACAGCCGAGCCCGGGGCTGCGCCAGGCGCTGACAGCCCT 180
DB 121 CTCCCTGTACCGCTCTCCACAGCCGAGCCCGGGGCTGCGCCAGGCGCTGACAGCCCT 180

QY 181 GGCTCTCTGATGCCCCCAAGCTCCCTCTCTGAGAGCCACAGCACCACCCAGACTGG 240
DB 181 GGCTCTCTGATGCCCCCAAGCTCCCTCTCTGAGAGCCACAGCACCACCCAGACTGG 240

QY 241 GGCAGCGCCGAGGACGAGCTGGGCCAGTGGGCCAGAGCGCCGAGGCGCGGCGG 300
DB 241 GGCAGCGCCGAGGACGAGCTGGGCCAGTGGGCCAGAGCGCCGAGGCGCGGCGG 300

QY 301 CCACATGGCCCAAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 CCACATGGCCCAAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 361 CCCACGACCCAGCAGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 CCCACGACCCAGCAGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 421 TGGGGCTGCGGCTGCGCCGGGAGACCGAGCAGAGCCCGAGGAGCCCGCGGAGGGGCA 480
DB 421 TGGGGCTGCGGCTGCGCCGGGAGACCGAGCAGAGCCCGAGGAGCCCGCGGAGGGGCA 480

QY 481 GCTTTGTGGAGATGTTGGACAACTTGAAGGGCAAGTTCGGGGGAGGGGCTACTACGTG 540
DB 481 GCTTTGTGGAGATGTTGGACAACTTGAAGGGCAAGTTCGGGGGAGGGGCTACTACGTG 540

QY 541 TGACCGTGGGAGCCCGCGGAGAGCGCTCAACATCTGTGGATACAGGAGCAGTAACT 600
DB 541 TGACCGTGGGAGCCCGCGGAGAGCGCTCAACATCTGTGGATACAGGAGCAGTAACT 600

QY 601 TTGCAGTGGGTGCTGCCCCCACCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 TTGCAGTGGGTGCTGCCCCCACCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 1741 AGTGGCGCTGCCCTCCGCTGCTGCGCCAGCAGCATGATGATCTTGTGTGATGACATCTCC 1800

QY 661 GCACATACCGGAGCTCCGGAAGGCTGTGTATGTGCCCTACACCCAGGCGAAGTGGGAAG 720
DB 661 GCACATACCGGAGCTCCGGAAGGCTGTGTATGTGCCCTACACCCAGGCGAAGTGGGAAG 720

QY 721 GGGAGCTGGGACCGACCTGTAGCATCCGCCATGGCCCCAACGTCACCTGTGGTGGCA 780
DB 721 GGGAGCTGGGACCGACCTGTAGCATCCGCCATGGCCCCAACGTCACCTGTGGTGGCA 780

QY 781 ACATTGCTGCGCATCAGCAAGTCTTTCATCAACGCTCCCACTGGGAAGCA 840
DB 781 ACATTGCTGCGCATCAGCAAGTCTTTCATCAACGCTCCCACTGGGAAGCA 840

QY 841 TCCTGGGCTGGCCCTATGCTGAGATTGCCAGCCCTGACGACTCCCTGGAGCCCTTCTTTG 900
DB 841 TCCTGGGCTGGCCCTATGCTGAGATTGCCAGCCCTGACGACTCCCTGGAGCCCTTCTTTG 900

QY 901 ACTCTCTGTTAAACGACGACCCAGCTTCCCAACCTCTTCTCCCTGCGAGCTTGTGGTCTG 960
DB 901 ACTCTCTGTTAAACGACGACCCAGCTTCCCAACCTCTTCTCCCTGCGAGCTTGTGGTCTG 960

QY 961 GCTTCCCTCTCAACGAGTCTGAAGTCTGCTGCGGAGGAGCATCATCTTGAG 1020
DB 961 GCTTCCCTCTCAACGAGTCTGAAGTCTGCTGCGGAGGAGCATCATCTTGAG 1020

QY 1021 GTATCGACCTGCTGTACACAGGAGCTCTCTGTGTATACACCCATCCGCGGAGTGGT 1080
DB 1021 GTATCGACCTGCTGTACACAGGAGCTCTCTGTGTATACACCCATCCGCGGAGTGGT 1080

QY 1081 ATTATGAGTGTATCTTCTGCGGCTGGAGATCAATGGACAGGATCTGAAATGGACTGCA 1140
DB 1081 ATTATGAGTGTATCTTCTGCGGCTGGAGATCAATGGACAGGATCTGAAATGGACTGCA 1140

QY 1141 AGGAGTACAACTATGACAAGAGCATTTGTGGACAGTGGCACCAACCACTTCGTTTGGCCA 1200
DB 1141 AGGAGTACAACTATGACAAGAGCATTTGTGGACAGTGGCACCAACCACTTCGTTTGGCCA 1200

QY 1201 AGAAGTGTTTGAAGCTGCAATCCATCAAGGAGCCCTCTCCACGAGGAAGTTC 1260
DB 1201 AGAAGTGTTTGAAGCTGCAATCCATCAAGGAGCCCTCTCCACGAGGAAGTTC 1260

QY 1261 CTGATGTTCTGCTAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 CTGATGTTCTGCTAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320

QY 1321 ACATTTTCCAGTCTCTCACTCTACCTTAATGGGTGAGTTTACCAACCACTTCCTTCGCGCA 1380
DB 1321 ACATTTTCCAGTCTCTCACTCTACCTTAATGGGTGAGTTTACCAACCACTTCCTTCGCGCA 1380

QY 1381 TCACCATCTCTCCGAGCAATACCTGCGGCGAGTGGAGATGTGGCCACGTCACAGAGC 1440
DB 1381 TCACCATCTCTCCGAGCAATACCTGCGGCGAGTGGAGATGTGGCCACGTCACAGAGC 1440

QY 1441 ACTGTTACAAAGTTTGGCATCTCACAGTCTATCCACGGGACTGTTATGGGAGCTGTTATCA 1500
DB 1441 ACTGTTACAAAGTTTGGCATCTCACAGTCTATCCACGGGACTGTTATGGGAGCTGTTATCA 1500

QY 1501 TGGAGGCTTCTAGCTTGTCTTTGATCGGGCCGAAAGAAATTTGCTTGTCTGCTACAGC 1560
DB 1501 TGGAGGCTTCTAGCTTGTCTTTGATCGGGCCGAAAGAAATTTGCTTGTCTGCTACAGC 1560

QY 1561 CTTTGGCATGTGACGATGAGTTTACGAGCGGAGCGGTGGAAGCCCTTTTGTTCACCTTGG 1620
DB 1561 CTTTGGCATGTGACGATGAGTTTACGAGCGGAGCGGTGGAAGCCCTTTTGTTCACCTTGG 1620

QY 1621 ACATGGAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG 1680
DB 1621 ACATGGAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG 1680

QY 1681 CCTATGTCATGGCTGCGCCCTCTTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 1681 CCTATGTCATGGCTGCGCCCTCTTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

QY 1741 AGTGGCGCTGCCCTCCGCTGCTGCGCCAGCAGCATGATGATCTTGTGTGATGACATCTCC 1800

PF 17-JUL-2001; 2001WO-1L00653.
XX
XX PR 18-JUL-2000; 2000IL-0137345.
XX PR 15-DEC-2000; 2000IL-0140354.
XX PA (COMP-) COMPUGEN LTD.
XX
XX PI Mintz L, Freilich S, Bernstein J;
XX DR WPI: 2002-155037/20.
XX DR P-PSDB; AB06120.
XX
XX One hundred and twenty eight novel nucleic acid sequences, useful for
XX treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
XX Claim 1; Page 124-126; 290pp; English.
XX
XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
XX encoding the proteins given in AB06037 to AB06164. The novel sequences
XX (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
XX antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
XX vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
XX anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
XX anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiac,
XX immunomodulator, anticonvulsant, antidiabetic, tranquiliser, antiulcer,
XX antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
XX nootropic and contraceptive activities. The NS can be used in vaccines,
XX gene therapy and antisense therapy. Nucleic acids, expression vectors and
XX antibodies from the present invention can be used for treating and
XX diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
XX diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
XX cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
XX glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
XX disease, coagulation disease, ischaemia, hypertension, asthma, immune
XX disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
XX depression, schizophrenia, viral disease, gastric ulcers, stroke,
XX Alzheimer's disease and as a contraceptive.
XX
XX Sequence 5757 BP; 1485 A; 1490 C; 1392 G; 1388 T; 2 other;

Query Match 92.9%; Score 2181.2; DB 24; Length 5757;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 2268; Conservative 1; Mismatches 4; Indels 75; Gaps 1;

QY 1 CCATGCGGCGCTCACAGCCCGCGGAGCCCGAGCCGCTGCCAGGCTGGCGCGG 60
DB 150 CCATGCGGCGCTCCAGCGCCCGCGGAGCCCGCGGCTGCCAGGCTGGCGCGG 209
QY 61 CSGTGCGGATGTAGCGGCTCCGGATCCCGAGCTCTCCCTGCTCCCGTCTGCGGAT 120
DB 210 CCGTGCGGATGTAGCGGCTCCGGATCCCGAGCTCTCCCTGCTCCCGTCTGCGGAT 269
QY 121 CTCCTGACCGCTCCACAGCCCGAGCCCGGGGCTGCCAGGCGCTGCGAGGCCCT 180
DB 270 CTCCTGACCGCTCTCCACAGCCCGAGCCCGGGGCTGGCCAGGCGCTGCGAGGCCCT 329
QY 181 GCGGCTCTGATGCCCGCCAGCTCCCTCTCTCTGAGAGCCACACAGCACCACAGACTTG 240
DB 330 GCGGCTCTGATGCCCGCCAGCTCCCTCTCTCTGAGAGCCACACAGCACCACAGACTTG 389
QY 241 GGCAGCGCCGAGGAGGAGCTGGGCGAGTGAGGCCCGCAGAGGCCCGGAGCGCGGGC 300
DB 390 GGCAGCGCCGAGGAGGAGCTGGGCGAGTGAGGCCCGCAGAGGCCCGGAGCGCGGGC 449
QY 301 CCACCATGGCCCAAGCCCTGCCCTGCTCTCTGCTGGATGGCGCGGGAGTCTGCCTG 360
DB 450 CCACCATGGCCCAAGCCCTGCCCTGCTCTCTGCTGGATGGCGCGGGAGTCTGCCTG 509
QY 361 CCCACGGCACCACAGCAGCGGATCCGGCTGCCCTGCGCAGCGGCTGGGGGGCGCCCCC 420
DB 510 CCCACGGCACCACAGCAGCGGATCCGGCTGCCCTGCGCAGCGGCTGGGGGGCGCCCCC 569
QY 421 TGGGGGTGGGGCTGCCCCCGGGAGACCCAGAGAGCCCGAGAGCCCGCGCGGAGGGCA 480

DB 570 TGGGGGTGGGGTGGCCCGGAGACCGAGAGCCCGAGAGCCCGCGCGGAGGGGCA 629
QY 481 GCTTTGTGGAGATGTGGACAACCTGAGGGGCAAGTGGGGGAGGGCTACTAGCTGGAGA 540
DB 630 GCTTTGTGGAGATGTGGACAACCTGAGGGGCAAGTGGGGGAGGGCTACTAGCTGGAGA 689
QY 541 TGACCGTGGGCGAGCCCGCGCAGACGCTCAACATCTCTGTGGATACAGGCGAGTAACCT 600
DB 690 TGACCGTGGGCGAGCCCGCGCAGACGCTCAACATCTCTGTGGATACAGGCGAGTAACCT 749
QY 601 TTGACGTGGGTGCTGCCCCCGCCACCTTCTTGCATCTGCTACTACCGAGGAGCTGTCCA 660
DB 750 TTGACGTGGGTGCTGCCCCCGCCACCTTCTTGCATCTGCTACTACCGAGGCGAGCTGTCCA 809
QY 661 GCACATACCGGGGACCTCCCGAAGGCTGTATGTGCCCTACACCGAGGGCAAGTGGGAAG 720
DB 810 GCACATACCGGGGACCTCCCGAAGGCTGTATGTGCCCTACACCGAGGGCAAGTGGGAAG 869
QY 721 GGGAGCTGGGCGACCGGACCTGTAAGCATCCCATGGCCCAAGCTCACCTGCTGCTGCCA 780
DB 870 GGGAGCTGGGCGACCGGACCTGTAAGCATCCCATGGCCCAAGCTCACCTGCTGCTGCCA 929
QY 781 ACATTTCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCA 840
DB 930 ACATTTCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCA 989
QY 841 TCTTGGGGTGGGCTATGCTGAGATTTGCCAGGCTGACGACTCCCTGGAGGCCCTTTCTTTG 900
DB 990 TCTTGGGGTGGGCTATGCTGAGATTTGCCAGGCTGACGACTCCCTGGAGGCCCTTTCTTTG 1020
QY 901 ACTTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTCTCCCTGCGAGCTTTGGTGGCTG 960
DB 1021 -----GCTTTGGTGGCTG 1034
QY 961 GCTTCCCTCCCAACAGTCTGAAGTGTGGCTCTGTGGAGGAGGAGCATGATCATTTGGAG 1020
DB 1035 GCTTCCCTCCCAACAGTCTGAAGTGTGGCTCTGTGGAGGAGGAGCATGATCATTTGGAG 1094
QY 1021 GTATCCAGCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCGCGGGAGTGGT 1080
DB 1095 GTATCCAGCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCGCGGGAGTGGT 1154
QY 1081 ATTATGAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGAGCTGCA 1140
DB 1155 ATTATGAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGAGCTGCA 1214
QY 1141 AGGAGTACAACTATGACAAGAGCATTTGTGGACAGTGGCACACCAACCTTTGTTGGCCA 1200
DB 1215 AGGAGTACAACTATGACAAGAGCATTTGTGGACAGTGGCACACCAACCTTTGTTGGCCA 1274
QY 1201 AGAAGTGTGTTGAGCTGAGTCAAAATCCATCAAGCAGGCTCTCCACCGAGGAAGTTCC 1260
DB 1275 AGAAGTGTGTTGAGCTGAGTCAAAATCCATCAAGCAGGCTCTCCACCGAGGAAGTTCC 1334
QY 1261 CTGATGTTTCTGGCTAGGAGAGCAGTGGTGTCTGGCAAGGAGGACCAACCTTTGTTGGA 1320
DB 1335 CTGATGTTTCTGGCTAGGAGAGCAGTGGTGTCTGGCAAGGAGGACCAACCTTTGTTGGA 1394
QY 1321 ACATTTTCCAGTCACTCTACTCTACCTAATGGTGGAGTTACCAACAGTCTCTTCGCA 1380
DB 1395 ACATTTTCCAGTCACTCTACTCTACCTAATGGTGGAGTTACCAACAGTCTCTTCGCA 1454
QY 1381 TCACCATCTTCCGAGCAATACCTGGGCGAGTGGGAAGATGTGGCCACGCTCCCAAGACG 1440
DB 1455 TCACCATCTTCCGAGCAATACCTGGGCGAGTGGGAAGATGTGGCCACGCTCCCAAGACG 1514
QY 1441 ACTGTTACAAAGTTTGGCATCTCAGTCACTCCAGGCGACTGTTATGGAGCTGTTATCA 1500
DB 1515 ACTGTTACAAAGTTTGGCATCTCAGTCACTCCAGGCGACTGTTATGGAGCTGTTATCA 1574
QY 1501 TGGAGGCGCTTCTACGTTGCTTTTGTATCGGCGCCGAAACGAATTTGCTGTCTACGCG 1560

Db	1575	TGGAGGCTTCTACGTTGCTTTTGATGGGCCGGAACGAATTGCTTTGCTGTGACGG	1634
Qy	1561	CTTGCCATGTGCACGATGAGTTTCAGGACGGCAGCGTGGAAAGCCCTTTTGTACACCTTGG	1620
Db	1635	CTTGCCATCTGCCACGATGAGTTTCAGGACGGCAGCGTGGAAAGCCCTTTTGTACACCTTGG	1694
Qy	1621	ACATGGGAAGACTGTGGCTACAACATTCCACACACAGATGAGTCAACCCCTCATGACCATAG	1680
Db	1695	ACATGGGAAGACTGTGGCTACAACATTTCACACACAGATGAGTCAACCCCTCATGACCATAG	1754
Qy	1681	CCTATGTCTATGGCTGCCATCTGCGGCCCTTTCATGCTGCGCACTCTGCCCTCATGGTGTGC	1740
Db	1755	CCTATGTCTATGGCTGCCATCTGCGGCCCTTTCATGCTGCGCACTCTGCCCTCATGGTGTGC	1814
Qy	1741	AGTGGCGTGCCTCCGCTGCCCTGCGCCAGCAGCATGATGACTTTTGTGTATGACATCTCCC	1800
Db	1815	AGTGGCGTGCCTCCGCTGCCCTGCGCCAGCAGCATGATGACTTTTGTGTATGACATCTCCC	1874
Qy	1801	TGCTGAAGTGAAGAGGCCCATGGGCAGAGATTAGAGATTCCCTTGACACACACCTCCGTG	1860
Db	1875	TGCTGAAGTGAAGAGGCCCATGGGCAGAGATTAGAGATTCCCTTGACACACACCTCCGTG	1934
Qy	1861	GTTCACTTTGGTCAACAAGTGAAGAGACACAGATGGCACTGTGGCCAGAGCACCTCAGGAC	1920
Db	1935	GTTCACTTTGGTCAACAAGTGAAGAGACACAGATGGCACTGTGGCCAGAGCACCTCAGGAC	1994
Qy	1921	CCTCCCCACCCACCAATGCCCTCTGCCCTTGATGGAGAAAGGCTGGCAAGGTGGGT	1980
Db	1995	CCTCCCCACCCACCAATGCCCTCTGCCCTTGATGGAGAAAGGCTGGCAAGGTGGGT	2054
Qy	1981	TCCAGGGACTGTACCTGTAGGAACACAGAAACAGAGAAAGAGACACTCTGCTGGCGGGA	2040
Db	2055	TCCAGGGACTGTACCTGTAGGAACACAGAAACAGAGAAAGAGACACTCTGCTGGCGGGA	2114
Qy	2041	ATACTCTTTGGTCACTCAAAATTTAAGTGGGAAATTCCTGCTGCTGAAACTTCAGCCCTG	2100
Db	2115	ATACTCTTTGGTCACTCAAAATTTAAGTGGGAAATTCCTGCTGCTGAAACTTCAGCCCTG	2174
Qy	2101	AACCTTTCTGCACCATTCCTTTAAATTCCCAACCCAAAGTATTCCTCTTTTCTTAGTTT	2160
Db	2175	AACCTTTCTGCACCATTCCTTTAAATTCCCAACCCAAAGTATTCCTCTTTTCTTAGTTT	2234
Qy	2161	CAGAAGTACTGGCATCACACGAGGTTACCTTTGGCGTGTCCCTGTGTTACCCCTGGCAG	2220
Db	2235	CAGAAGTACTGGCATCACACGAGGTTACCTTTGGCGTGTGTCCCTGTGTTACCCCTGGCAG	2294
Qy	2221	AGAAGAGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGACACAGTTTG	2280
Db	2295	AGAAGAGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGACACAGTTTG	2354
Qy	2281	CTATTTGCTTTTAGACACAGGGACTGTATAACAAGGCTTAACATTGCTGCAAAAGATTGCCCT	2340
Db	2355	CTATTTGCTTTTAGACACAGGGACTGTATAACAAGGCTTAACATTGCTGCAAAAGATTGCCCT	2414
Qy	2341	CTTGAATT	2348
Db	2415	CTTGAATT	2422

RESULT 4

RES001.4	
AAAL5662	
ID	AAAL5662 standard; cDNA; 2070 BP.
XX	
XX	
AC	AAAL5662;
XX	
DT	03-AUG-2000 (first entry)
XX	
DE	Human aspartyl protease 2 (a) (Asp2) nucleotide sequence.
XX	
XX	
KW	Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 2;
KW	Alzheimer's disease; beta secretase site; ss.
XX	
OS	Homo sapiens.

QY 726 CTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCAAGCTCACTGTGGTGCCAACTT 785
DB 421 CTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCAAGCTCACTGTGGTGCCAACTT 480
QY 786 GCTGGCATCACTGAATCAGACAAAGTTCTTCATCAACGGGTCCAACTGGGAAGGCATCTCTG 845
DB 481 GCTGGCATCACTGAATCAGACAAAGTTCTTCATCAACGGGTCCAACTGGGAAGGCATCTCTG 540
QY 846 GGGCTGGCTATTGCTGAGATGCGAGGCTGAGCATCTCCCTGGAGCTTTCTTTGACTCT 905
DB 541 GGGCTGGCTATTGCTGAGATGCGAGGCTGAGCATCTCCCTGGAGCTTTCTTTGACTCT 600
QY 906 CTGGTAAGCAGACCCACCGTTCCTCAACCTCTCTCCCTGCAGCTTTGTTGGTGTGCTTC 965
DB 601 CTGGTAAGCAGACCCACCGTTCCTCAACCTCTCTCCCTGCAGCTTTGTTGGTGTGCTTC 660
QY 966 CCGCTCAACGAGTCTGAAGTGTGGCTCTGTGCGGAGGAGCATGATCATTTGGAGGTATC 1025
DB 661 CCGCTCAACGAGTCTGAAGTGTGGCTCTGTGCGGAGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACTCGCTGACACAGCAGTCTCTGTATACACCCATCGCGGGAGTGGTATTAT 1085
DB 721 GACCACTCGCTGACACAGCAGTCTCTGTATACACCCATCGCGGGAGTGGTATTAT 780
QY 1086 GAGGTGATCATTTGTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 1145
DB 781 GAGGTGATCATTTGTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 840
QY 1146 TACAACATGACAAAGCATTTGGAGACATGGGACAGTGGACACCAACCTTCGTTTGGCCCAAGAA 1205
DB 841 TACAACATGACAAAGCATTTGGAGACATGGGACAGTGGGACCAACCAACCTTCGTTTGGCCCAAGAA 900
QY 1206 GTCTTTGAAGCTGCACTCAATTCATCAAGGCAGCTCTCCACGGAGAGTTCCCTGAT 1265
DB 901 GTCTTTGAAGCTGCACTCAATTCATCAAGGCAGCTCTCCACGGAGAGTTCCCTGAT 960
QY 1266 GGTCTTGGCTAGGAGCAGCTGGTGTGCTGGCAAGCAGGACACCCCTTGGAAACATT 1325
DB 961 GGTCTTGGCTAGGAGCAGCTGGTGTGCTGGCAAGCAGGACACCCCTTGGAAACATT 1020
QY 1326 TTCCCGAGTCACTCACTCTACCTAATGGGTGAGGTTACCAACAGTCTTCCCGCATCAC 1385
DB 1021 TTCCCGAGTCACTCACTCTACCTAATGGGTGAGGTTACCAACAGTCTTCCCGCATCAC 1080
QY 1386 ATCTTTCCGAGCAATACCTCGCGGCAGTGGGAAGATGTGGCCACGTCCTCCCAAGACGACTGT 1445
DB 1081 ATCTTTCCGAGCAATACCTCGCGGCAGTGGGAAGATGTGGCCACGTCCTCCCAAGACGACTGT 1140
QY 1446 TACAAGTTTGGCATCTCACAGTCAATCCACGGGCATCTTTATGGGAGCTGTTATCATGGAG 1505
DB 1141 TACAAGTTTGGCATCTCACAGTCAATCCACGGGCATCTTTATGGGAGCTGTTATCATGGAG 1200
QY 1506 GGTCTTACGTTGTCTTTGATCGGGCCGGAACGAAATGGCTTTGCTGTACGCGCTTGC 1565
DB 1201 GGTCTTACGTTGTCTTTGATCGGGCCGGAACGAAATGGCTTTGCTGTACGCGCTTGC 1260
QY 1566 CATGTGCAGATCAGTTTCAGGACGGCAGCGGTGGAGGCCCTTTTGTCACTTTGGACATG 1625
DB 1261 CATGTGCAGATCAGTTTCAGGACGGCAGCGGTGGAGGCCCTTTTGTCACTTTGGACATG 1320
QY 1626 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
DB 1321 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380
QY 1686 GTCATGGCTGCCATCTGCGGCCCTTCTCATGCTGCCACTCTGCCCTCATGTTGTGTGAGTGG 1745
DB 1381 GTCATGGCTGCCATCTGCGGCCCTTCTCATGCTGCCACTCTGCCCTCATGTTGTGTGAGTGG 1440
QY 1746 CGTGTGCTCGCTGCTGCGCCAGCAGCATGATGACTTTTCTGATGACATCTCCCTGCTG 1805
DB 1441 CGTGTGCTCGCTGCTGCGCCAGCAGCATGATGACTTTTCTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGAGGAGGCCCATGGGGCAGAAGATAGAGATTCCCTTGGACACACCTCCGTGGTTCA 1865

DB 1501 AAGTGAGGAGGCCCATGGGGCAGAAGATAGAGATTCCCTTGGACACACCTCCCGTGGTTCA 1560
QY 1866 CTTTGTGTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCAGCTCAGACCCCTCC 1925
DB 1561 CTTTGTGTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCAGCTCAGACCCCTCC 1620
QY 1926 CACCCACCAAAATGCCCTTGCCTTGTATGGAGAAGGAAAGCCTGGCAAGGTGGTTCCAG 1985
DB 1621 CACCCACCAAAATGCCCTTGCCTTGTATGGAGAAGGAAAGCCTGGCAAGGTGGTTCCAG 1680
QY 1986 GGAAGTGTACTGTAGGAAACAGAAAAGAGAAGAAAGCACTCTGCTGGCGGAATACT 2045
DB 1681 GGAAGTGTACTGTAGGAAACAGAAAAGAGAAGAAAGCACTCTGCTGGCGGAATACT 1740
QY 2046 CTTTGGTCACTCAAAATTAAGTCGGGAATTTCTGCTGCTTGAACCTTCAGCCCTGACCT 2105
DB 1741 CTTTGGTCACTCAAAATTAAGTCGGGAATTTCTGCTGCTTGAACCTTCAGCCCTGACCT 1800
QY 2106 TTTCTCCACCAATTCCTTTAAATTTCTCAACCCAAAGATTTCTTTTCTTTAGTTTCAGAA 2165
DB 1801 TTTCTCCACCAATTCCTTTAAATTTCTCAACCCAAAGATTTCTTTTCTTTAGTTTCAGAA 1860
QY 2166 GTACTGGCATCACGAGGTTACCTTTGGCGTGTGCTCCCTGTGTTGCTGCTGGCAGAGAAG 2225
DB 1861 GTACTGGCATCACGAGGTTACCTTTGGCGTGTGCTCCCTGTGTTGCTGCTGGCAGAGAAG 1920
QY 2226 AGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCACTAGGAGAGGATGCACAGTTTGCTATT 2285
DB 1921 AGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCACTAGGAGAGGATGCACAGTTTGCTATT 1980
QY 2286 TGTCTTAGACACAGGACTGTATAAACAGCCTAAACATTTGTTGCAAAAGATTGCTCTTGA 2345
DB 1981 TGTCTTAGACACAGGACTGTATAAACAGCCTAAACATTTGTTGCAAAAGATTGCTCTTGA 2040
QY 2346 ATT 2348
DB 2041 ATT 2043
RESULT 5
AADI7865
ID 'AADI7865 standard; cDNA; 2070' BP.
XX
AC AADI7865;
XX
DT 10-DEC-2001 (first entry)
XX
Human aspartyl protease 2(a) [hu-Asp2(a)] cDNA.
DE
KW Human; aspartyl protease 2(a); Asp2(a); amyloid precursor protein; APP;
KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
KW amyloid plaque; neuronal loss; proteolytic; nontropic; neuroprotective;
KW chromosome 11q23.3-24.1; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH CDS 1..1506
FT /*tag= a
FT /product= "Human aspartyl protease 2(a)"
FT sig_peptide 1..63
FT /*tag= b
FT mat_peptide 64..1503
FT /*tag= c
FT /product= "Mature human aspartyl protease 2(a)"
PN GB2357767-A.
XX
PD 04-JUL-2001.
XX
PF 22-SEP-2000; 2000GB-0023315.
XX

Db 421 CTGGGACCCAGCCGGTAAGCATCCCCATGGCCCCAAGCTCACTGTGGGTGCCAACATT 480
Qy 786 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGAAGGCATCCCTG 845
Db 481 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGAAGGCATCCCTG 540
Qy 846 GGCTTGGCCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCCTTTGACTCT 905
Db 541 GGCTTGGCCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCCTTTGACTCT 600
Qy 906 CTGGTAAAGCAGACCCAGCTTCCAAACCTCTTCTCCCTGCGACGCTTGTGTGGTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCAGCTTCCAAACCTCTTCTCCCTGCGACGCTTGTGTGGTGGCTTC 660
Qy 966 CCCCTCAACCACTGCTGAAGTCTGGCCCTGTGCGAGGAGGATGATCATTTGGAGTATC 1025
Db 661 CCCCTCAACCACTGCTGAAGTCTGGCCCTGTGCGAGGAGGATGATCATTTGGAGTATC 720
Qy 1026 GACCACTCGCTGTACACAGGAGCTCTCTGGTATACACCCATCCGGGGGAGTGGTATTAT 1085
Db 721 GACCACTCGCTGTACACAGGAGCTCTCTGGTATACACCCATCCGGGGGAGTGGTATTAT 780
Qy 1086 GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCCAAGGAG 1145
Db 781 GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCCAAGGAG 840
Qy 1146 TACAACCTATGACAAGAGCATTTGGACAGTGGCACACCAACCTTTCCTTGGCCCAAGAAA 1205
Db 841 TACAACCTATGACAAGAGCATTTGGACAGTGGCACACCAACCTTTCCTTGGCCCAAGAAA 900
Qy 1206 GTGTTTGAAGCTGCAAGTCAATTCATCAAGGAGCCTCTCCACGGAGAGTTCCCTGAT 1265
Db 901 GTGTTTGAAGCTGCAAGTCAATTCATCAAGGAGCCTCTCCACGGAGAGTTCCCTGAT 960
Qy 1266 GGTTCCTGGCTAGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCAACCCCTTGGACATT 1325
Db 961 GGTTCCTGGCTAGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCAACCCCTTGGACATT 1020
Qy 1326 TTCCCACTCATCTCACCTACTAATGGGTGAGGTTACCAACAGTCCCTTCCGCATCACC 1385
Db 1021 TTCCCACTCATCTCACCTACTAATGGGTGAGGTTACCAACAGTCCCTTCCGCATCACC 1080
Qy 1386 ATCCTTCCGAGCAATACCTCGGCGCAGTGGAAAGATGTGGCCACAGTCCCAAGACGACTGT 1445
Db 1081 ATCCTTCCGAGCAATACCTCGGCGCAGTGGAAAGATGTGGCCACAGTCCCAAGACGACTGT 1140
Qy 1446 TACAAGTTTGCCTATCTCACAGTCAATCCACGGCAGCTGTATGGAGCTCTTATCATGGAG 1505
Db 1141 TACAAGTTTGCCTATCTCACAGTCAATCCACGGCAGCTGTATGGAGCTCTTATCATGGAG 1200
Qy 1506 GGCTTCTACGTTGCTTTGATCGGGCCCGGAAACGAAATTTGGCTTTGCTGTCAGCGCTTGC 1565
Db 1201 GGCTTCTACGTTGCTTTGATCGGGCCCGGAAACGAAATTTGGCTTTGCTGTCAGCGCTTGC 1260
Qy 1566 CATGTGCAGCATGAGTTTCAGACGGCAGCGGTGGAAAGGCCCTTTTGTACCTTTGGACATG 1625
Db 1261 CATGTGCAGCATGAGTTTCAGACGGCAGCGGTGGAAAGGCCCTTTTGTACCTTTGGACATG 1320
Qy 1626 GAAGACTGTGGCTACACATTCCACAGACAGATGATGATCAACCCCTCATGACCATAGCCTAT 1685
Db 1321 GAAGACTGTGGCTACACATTCCACAGACAGATGATGATCAACCCCTCATGACCATAGCCTAT 1380
Qy 1686 GTCATGGCTGCCATCTGGCCCTCTTCATGTGCGACCTCTGCTCATGTGTGTCAAGTGG 1745
Db 1381 GTCATGGCTGCCATCTGGCCCTCTTCATGTGCGACCTCTGCTCATGTGTGTCAAGTGG 1440
Qy 1746 CGCTGCTCCGCTGCCCTGGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGCTCCGCTGCCCTGGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1500
Qy 1806 AAGTGGAGGCCCCATGGGAGAGATAGATTTCCCTTGGACCAACCTCCGTGGTTCA 1865
|||||

Db 1501 AAGTCAGGAGGCCCATGGGAGAGATAGAGATTCCCTTGACCACACCTCCGTGTTCA 1560
Qy 1866 CTTTGGTTCACAAGTAGGAGACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCTCC 1925
Db 1561 CTTTGGTTCACAAGTAGGAGACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCTCC 1620
Qy 1926 CCACCCACCAAAATGCTCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGGTGGTTCAG 1985
Db 1621 CCACCCACCAAAATGCTCTCTGCTTGTATGGAGAGAAAGGCTGGCAAGGTGGTTCAG 1680
Qy 1986 GGACTGTACTGTAGGAAACAGAGAAAGAGACACTCTGCTGGGGGAATACT 2045
Db 1681 GGACTGTACTGTAGGAAACAGAGAAAGAGACACTCTGCTGGGGGAATACT 1740
Qy 2046 CTTTGGTTCACCTCAAAATTAAGTCGGGAAATTCGCTCTTCAAACTTCAGCCCTCAACT 2105
Db 1741 CTTTGGTTCACCTCAAAATTAAGTCGGGAAATTCGCTCTTCAAACTTCAGCCCTCAACT 1800
Qy 2106 TTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGTATTTCTTTCTTATAGTTTCAGAA 2165
Db 1801 TTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGTATTTCTTTCTTATAGTTTCAGAA 1860
Qy 2166 GTACTGGCATCACACCCAGGTTTACCTTGGCGTGTGTCCTGTTGGTACCTGGCAGAGAG 2225
Db 1861 GTACTGGCATCACACCCAGGTTTACCTTGGCGTGTGTCCTGTTGGTACCTGGCAGAGAG 1920
Qy 2226 AGACCAAGCTTGTTCCTGCTGGCCAAAGTTCAGTAGGAGAGATGCACAGTTTCTATT 2285
Db 1921 AGACCAAGCTTGTTCCTGCTGGCCAAAGTTCAGTAGGAGAGATGCACAGTTTCTATT 1980
Qy 2286 TGCTTTAGAGACAGGAGCTGTATAACACAGCCCTAACATTTGGTGGAAAGATTGCCCTTTGA 2345
Db 1981 TGCTTTAGAGACAGGAGCTGTATAACACAGCCCTAACATTTGGTGGAAAGATTGCCCTTTGA 2040
Qy 2346 ATT 2348
Db 2041 ATT 2043
RESULT 7
AAS11517
ID AAS11517 standard; cDNA; 2070 BP.
XX AC AAS11517;
XX XX
DT 24-OCT-2001 (first entry)
XX XX
DE Human cDNA encoding Aspartyl protease 2(a), Asp2(a).
XX XX
KW Human; Aspartyl protease; Asp2(a); beta-secretase; neurotropic;
KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
KW amyloid-beta; Abeta; ss.
XX XX
OS Homo sapiens.
XX XX
FH Key Location/Qualifiers
FT CDS 1..1506
FT FT /*tag= a
FT FT /product= "Asp2(a)"
FT FT sig_peptide 1..63
FT FT /*tag= b
FT FT sig_peptide 64..135
FT FT /*tag= c
FT FT /*label= pre_pro_peptide 136..171
FT FT /*tag= d
FT FT /*label= pro_peptide 172..1503
FT FT mat_peptide
FT FT /*tag= e
FT FT /*label= Mature_Asp2(a)
XX WO200149098-A2.
PN
XX


```
QY 486 GTGGAGATGTTGCAACACCTGAGGGCAAGTCGGGGCAGGCTACTACGTGGAGATGACC 545
|
|
|
Db 181 GTGGAGATGTTGACAACTGAGGGCAAGTCGGGGCAGGCTACTACGTGGAGATGACC 240
|
|
|
QY 546 GTGGGAGCCCCCGGAGAGCGCTCAACATCCTGGTGGATACAGGACGACGTAACCTTGGCA 605
|
|
|
Db 241 GTGGGAGCCCCCGGAGAGCGCTCAACATCCTGGTGGATACAGGACGACGTAACCTTGGCA 300
|
|
|
QY 606 GTGGGTGCTGCCGCCCCACCCCTTCTCTGCGATCGCTACTACCAGAGGCGAGCTGCCAGCACA 665
|
|
|
Db 301 GTGGGTGCTGCCGCCCCACCCCTTCTCTGCGATCGCTACTACCAGAGGCGAGCTGCCAGCACA 360
|
|
|
QY 666 TACCGGAGACCTCCGGAAAGGTTGTATGTGCCCTTACACCCAGGCAAGTGGGAAGGGGAG 725
|
|
|
Db 361 TACCGGAGACCTCCGGAAAGGTTGTATGTGCCCTTACACCCAGGCAAGTGGGAAGGGGAG 420
|
|
|
QY 726 CTGGGACCGACCTGTGTAAGCATCCCCCATGGCCCCAACCTGCTGTCGCTGCCAACATTT 785
|
|
|
Db 421 CTGGGACCGACCTGTGTAAGCATCCCCCATGGCCCCAACCTGCTGTCGCTGCCAACATTT 480
|
|
|
QY 786 GCTGCCATCACTGAATCAGACAAGTTCTTTCATCAACGGCTCCAACTGGGAAGGCATCCTG 845
|
|
|
Db 481 GCTGCCATCACTGAATCAGACAAGTTCTTTCATCAACGGCTCCAACTGGGAAGGCATCCTG 540
|
|
|
QY 846 GGGCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 905
|
|
|
Db 541 GGGCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 600
|
|
|
QY 906 CTGGTAAGCAGACCCACGTTCCCAAGCTCTTCTCCCTGCGAGCTTTGTGTGCTGGCTTC 965
|
|
|
Db 601 CTGGTAAGCAGACCCACGTTCCCAAGCTCTTCTCCCTGCGAGCTTTGTGTGCTGGCTTC 660
|
|
|
QY 966 CCCCTCAACAGTCTGAAGTGTGGCTCTGTTCGGAGGGAGCATGATCATTTGAGGTATC 1025
|
|
|
Db 661 CCCCTCAACAGTCTGAAGTGTGGCTCTGTTCGGAGGGAGCATGATCATTTGAGGTATC 720
|
|
|
QY 1026 GACCACTCGCTTACACAGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT 1085
|
|
|
Db 721 GACCACTCGCTTACACAGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT 780
|
|
|
QY 1086 GAGGTGATCATTTGGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 1145
|
|
|
Db 781 GAGGTGATCATTTGGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 840
|
|
|
QY 1146 TACAACATATGACAGAGCATTTGGAGAGTGGCAGTGGCACCACCAACCTTCGTTTGGCCCAAGAAA 1205
|
|
|
Db 841 TACAACATATGACAGAGCATTTGGAGAGTGGCAGTGGCACCACCAACCTTCGTTTGGCCCAAGAAA 900
|
|
|
QY 1206 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCGACGCTTCCTCCAGGAGAAGTTCCTTGAT 1265
|
|
|
Db 901 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCGACGCTTCCTCCAGGAGAAGTTCCTTGAT 960
|
|
|
QY 1266 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGTCGGAAGCAGGACACCCCTTGGAAACATT 1325
|
|
|
Db 961 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGTCGGAAGCAGGACACCCCTTGGAAACATT 1020
|
|
|
QY 1326 TTCCCACTCATCTACCTTACCTAATGGTTCAGTTTACCACCACTGCTTCCGATCAAC 1385
|
|
|
Db 1021 TTCCCACTCATCTACCTTACCTAATGGTTCAGTTTACCACCACTGCTTCCGATCAAC 1080
|
|
|
QY 1386 ATCTCTCCGAGCAATACCTCCGCGCCAGTGGAGATGTGGCCAGCTCCCAAGACGACTGT 1445
|
|
|
Db 1081 ATCTCTCCGAGCAATACCTCCGCGCCAGTGGAGATGTGGCCAGCTCCCAAGACGACTGT 1140
|
|
|
QY 1446 TACAAGTTTGGCATCTCACAGTCATCCACGGGCACTGTTATGGAGCTGTATCATCGGAG 1505
|
|
|
Db 1141 TACAAGTTTGGCATCTCACAGTCATCCACGGGCACTGTTATGGAGCTGTATCATCGGAG 1200
|
|
|
QY 1506 GGGTTCACGTTTGTCTTTGATCGGGCCCCGAAAACGAATTTGGCTTTGCTGTCAGGCGCTTC 1565
|
|
|
Db 1201 GGGTTCACGTTTGTCTTTGATCGGGCCCCGAAAACGAATTTGGCTTTGCTGTCAGGCGCTTC 1260
|
|
|
```

```
QY 1566 CATGTGCACGATGAGTTTCAGGAGGGCAGCGGTGGAAGGCCCTTTTGTACCTTTGCACATG 1625
|
|
|
Db 1261 CATGTGCACGATGAGTTTCAGGAGGGCAGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1320
|
|
|
QY 1626 GAGACTGTGGCTACAAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
|
|
|
Db 1321 GAGACTGTGGCTACAAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380
|
|
|
QY 1686 GTCATGCTGCCATCTCGGCCCTTTCATGCTGCCACTCTGCCCTCATGCTGTCAGTGG 1745
|
|
|
Db 1381 GTCATGCTGCCATCTCGGCCCTTTCATGCTGCCACTCTGCCCTCATGCTGTCAGTGG 1440
|
|
|
QY 1746 CGTGTCCCTCGCTGCTCGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1805
|
|
|
Db 1441 CGTGTCCCTCGCTGCTCGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1500
|
|
|
QY 1806 AAGTGAGGAGGCCCATGGGAGAGAATAGAGATTTCCCTTGACACACACCTCCCGTGGTTCA 1865
|
|
|
Db 1501 AAGTGAGGAGGCCCATGGGAGAGAATAGAGATTTCCCTTGACACACACCTCCCGTGGTTCA 1560
|
|
|
QY 1866 CTTTGTGTCAAGTAPAGGAGACACAGATGGCACCTGTGCCAGAGCACCTCAGGACCCCTCC 1925
|
|
|
Db 1561 CTTTGTGTCAAGTAPAGGAGACACAGATGGCACCTGTGCCAGAGCACCTCAGGACCCCTCC 1620
|
|
|
QY 1926 CCACCCACCAATGCCCTCTGCTTGTATGGAGAGAAAAGCTGGCAAGTGGGTTCAG 1985
|
|
|
Db 1621 CCACCCACCAATGCCCTCTGCTTGTATGGAGAGAAAAGCTGGCAAGTGGGTTCAG 1680
|
|
|
QY 1986 GGACTGTACCTGTAGGAAACAGAAAGAGAGAAAGAACGACTCTGTGGCGGGAATFACT 2045
|
|
|
Db 1681 GGACTGTACCTGTAGGAAACAGAAAGAGAGAAAGAACGACTCTGTGGCGGGAATFACT 1740
|
|
|
QY 2046 CTTTGTGCACCTCAAAATTTAAGTCGGGAAATTTCTGCTCTTCAAACTTTCAGCCCTGAACCT 2105
|
|
|
Db 1741 CTTTGTGCACCTCAAAATTTAAGTCGGGAAATTTCTGCTCTTCAAACTTTCAGCCCTGAACCT 1800
|
|
|
QY 2106 TTGTCCACCAATTCCTTTAAATTTCTCCAAACCCAAAGATTTCTCTTTTCTTAGTTTCAGAA 2165
|
|
|
Db 1801 TTGTCCACCAATTCCTTTAAATTTCTCCAAACCCAAAGATTTCTCTTTTCTTAGTTTCAGAA 1860
|
|
|
QY 2166 GTACTGCGATCACACGAGGTTACCTTTGGGCTGTGTCCTGTGTACCCCTGGGAGAGAG 2225
|
|
|
Db 1861 GTACTGCGATCACACGAGGTTACCTTTGGGCTGTGTCCTGTGTACCCCTGGGAGAGAG 1920
|
|
|
QY 2226 AGACCAAGCTGTGTTCCCTGCTGCCAAAGTCACTAGGAGAGGATGCACAGTTTGTCTATT 2285
|
|
|
Db 1921 AGACCAAGCTGTGTTCCCTGCTGCCAAAGTCACTAGGAGAGGATGCACAGTTTGTCTATT 1980
|
|
|
QY 2286 TGCTTTAGACAGAGGAGTGTATAAACAAGCCTAACATTTGGTGCAAGATTTGCCCTTTGA 2345
|
|
|
Db 1981 TGCTTTAGACAGAGGAGTGTATAAACAAGCCTAACATTTGGTGCAAGATTTGCCCTTTGA 2040
|
|
|
QY 2346 ATT 2348
|
|
|
Db 2041 ATT 2043
|
|
|
RESULT 9
AAD06739
ID AAD06739 standard; cdna; 2070 BP.
XX
AC AAD06739;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human aspartyl protease 2a (Asp2a) cdna.
XX
KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;
KW Alzheimer's disease; antialzheimer's; aspartyl protease 2a; Asp 2a;
KW beta-secretase; chromosome 11q23.3-24.1; ss.
XX
OS Homo sapiens.
XX
```

```

FH Key Location/Qualifiers
FT CDS 1..1506
FT FT /*tag= a
FT FT /product= "Human aspartyl protease 2a"
FT FT 1..63
FT FT /*tag= b
FT FT 64..135
FT FT /*tag= c
FT FT /note= "Pre-pro-peptide"
FT FT 136..171
FT FT /*tag= d
FT FT /note= "Pro-peptide"
FT FT 172..1503
FT FT /*tag= e
FT FT /product= "Human mature aspartyl protease 2a"
XX WO200123533-A2.
XX PD
XX 05-APR-2001.
XX XX
XX PF 22-SEP-2000; 2000WO-US26080.
XX PR 23-SEP-1999; 99US-0155493.
XX PR 23-SEP-1999; 99WO-US20881.
XX PR 13-OCT-1999; 99US-0416901.
XX PR 06-DEC-1999; 99US-0169232.
XX XX
XX PA (PHAA ) PHARMACIA & UPJOHN CO.
XX PI
XX Gurney M, Bienkowski MJ;
XX XX
XX WPI; 2001-290516/30.
XX DR P-PSDB; AAE02581.
XX PT
XX Enzymes that cleave the alpha-secretase site of the amyloid precursor
XX protein, useful for the treatment of Alzheimer's disease -
XX XX
XX Example 2; Page 126-127; 189pp; English.
XX CC The present invention relates to enzymes for cleaving the alpha-
XX secretase site of the amyloid precursor protein (APP) and methods of
XX identifying those enzymes. The methods may be used to identify enzymes
XX that may be used to cleave the alpha-secretase cleavage site of the APP
XX protein. The enzymes may be used to treat or modulate the progress of
XX Alzheimer's disease. The present sequence is human aspartyl protease
XX (Asp) 2a cDNA. Asp 2a has beta-secretase protease activity. Asp2 gene
XX is located on chromosome 11q23.3-24.1.
XX SQ Sequence 2070 BP; 476 A; 583 C; 562 G; 449 T; 0 other;

Query Match 86.9%; Score 2039.8; DB 22; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 306 ATGGCCCAAGCCCTGGCTCTGCTGTGGATGGCGCGGGAGTGTGCTGCCCTGCCAC 365
DB 1 ATGGCCCAAGCCCTGGCTCTGCTGTGGATGGCGCGGGAGTGTGCTGCCCTGCCAC 60
QY 366 GGCACCCAGCAGCGCATCCGCTGCCCTGCGCAGCGCTGGGGCGCCCTGGGG 425
DB 61 GGCACCCAGCAGCGCATCCGCTGCCCTGCGCAGCGCTGGGGCGCCCTGGGG 120
QY 426 CTGCGGCTGCCCGGGAGACCGAGACGCCAGAGCCCGGAGGCGCGGAGGGCAGCTTT 485
DB 121 CTCGCGCTGCCCGGGAGACCGAGACGCCAGAGCCCGGAGGCGCGGAGGGCAGCTTT 180
QY 486 GTGAGATGTTGGACAACTCGAGGGGCAAGTCGGGCGAGGGCTACTACGTGGAGATGACC 545
DB 181 GTGAGATGTTGGACAACTCGAGGGGCAAGTCGGGCGAGGGCTACTACGTGGAGATGACC 240
QY 546 GTGGGAGCCCCCGGAGACGCTCAACATCTGTTGGATACAGGAGAGTAGTAATTGCA 605
DB 241 GTGGGAGCCCCCGGAGACGCTCAACATCTGTTGGATACAGGAGAGTAGTAATTGCA 300

```

606 GTGGGTGCTGCCCCCACCCTTCTGTCATCGCTACTACAGAGGAGCTGTCCAGCACA 665
301 GTGGGTGCTGCCCCCACCCTTCTGTCATCGCTACTACAGAGGAGCTGTCCAGCACA 360
666 TACCGGAGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGAAGTGGGAAGGGAG 725
361 TACCGGAGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGAAGTGGGAAGGGAG 420
726 CTGGGACCCAGCTGGTAAGCATCCCCCATGGCCCCCAAGCTCACTGTGGGTGCCAACATT 785
421 CTGGGACCCAGCTGGTAAGCATCCCCCATGGCCCCCAAGCTCACTGTGGGTGCCAACATT 480
786 GGTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCTTG 845
481 GGTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCTTG 540
846 GGGCTGGCCTATGCTGAGATTTGCCAGGCTTGACGACTCTCCCTGGAGCTTTTCTTTGACTCT 905
541 GGGCTGGCCTATGCTGAGATTTGCCAGGCTTGACGACTCTCCCTGGAGCTTTTCTTTGACTCT 600
906 CTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGCACCTTTTGTGTGGCTGCTTC 965
601 CTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGCACCTTTTGTGTGGCTGCTTC 660
966 CCCTCAACCCAGTCTGAAGTCTGTGGCTCTGTCCGAGGAGCATGATCATTTGGAGGTATC 1025
661 CCCTCAACCCAGTCTGAAGTCTGTGGCTCTGTCCGAGGAGCATGATCATTTGGAGGTATC 720
1026 GACCACTGCTGTACAGAGCAGTCTCTGGTATACACCATCCGGCGGGAGTGGTATAT 1085
721 GACCACTGCTGTACAGAGCAGTCTCTGGTATACACCATCCGGCGGGAGTGGTATAT 780
1086 GAGGTATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 1145
781 GAGGTATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 840
1146 TACAACATATGACAAGAGCATTGTGGACAGTGGCACCACCAACTCTGTTTGCACAGAAA 1205
841 TACAACATATGACAAGAGCATTGTGGACAGTGGCACCACCAACTCTGTTTGCACAGAAA 900
1206 GTGTTTGAAGCTGAGTCAAAATCCATCAAGGAGCTCTCTCCAGGAGAAAGTTCCTGTAT 1265
901 GTGTTTGAAGCTGAGTCAAAATCCATCAAGGAGCTCTCTCCAGGAGAAAGTTCCTGTAT 960
1266 GGTTCCTGGCTAGGAGAGCAGCTGTGTGCTGGCAAGGAGGACACCCCTTCCGAACATT 1325
961 GGTTCCTGGCTAGGAGAGCAGCTGTGTGCTGGCAAGGAGGACACCCCTTCCGAACATT 1020
1326 TTCCCACTCATCTCACTCTACCTAATGGGTGAGGTTTACCAACAGTCTTTCGCATCAACC 1385
1021 TTCCCACTCATCTCACTCTACCTAATGGGTGAGGTTTACCAACAGTCTTTCGCATCAACC 1080
1386 ATCCCTTCCGAGCAATACCTGCGGCGAGTGGAAAGATGTGGCCACAGTCCCAAGAGCAGTGT 1445
1081 ATCCCTTCCGAGCAATACCTGCGGCGAGTGGAAAGATGTGGCCACAGTCCCAAGAGCAGTGT 1140
1446 TACAAGTTTGGCACTCTCACAGTCATCCACGGCAGCTGTATGGAGCTGTATCATGGAG 1505
1141 TACAAGTTTGGCACTCTCACAGTCATCCACGGCAGCTGTATGGAGCTGTATCATGGAG 1200
1506 GGTTCCTACGTTGTCTTTGTATGGGCGCCGAAACCAATTTGGCTGTCTGACGCGTTCG 1565
1201 GGTTCCTACGTTGTCTTTGTATGGGCGCCGAAACCAATTTGGCTGTCTGACGCGTTCG 1260
1566 CATGTCCAGCATGAGTTTCAGGAGCGGCTGGAAAGGCCCTTTTGTACCTTTGGACATG 1625
1261 CATGTCCAGCATGAGTTTCAGGAGCGGCTGGAAAGGCCCTTTTGTACCTTTGGACATG 1320
1626 GAAGACTGTGGCTTACAACATTTCCACAGACATGATGATCAACCTCATGACCATAGCCTAT 1685
1321 GAAGACTGTGGCTTACAACATTTCCACAGACATGATGATCAACCTCATGACCATAGCCTAT 1380

```
QY 1686 GTCATGGCTGCCATCTGCGCCCTTTCATGCTGCCACTCTGCCCTCATGTGTGTCTAGTGG 1745
DB 1381 GTCATGGCTGCCATCTGCGCCCTTTCATGCTGCCACTCTGCCCTCATGTGTGTCTAGTGG 1440
QY 1746 CGCTGCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1805
DB 1441 CGCTGCTCGCTGCGCTGCGCCAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTAGAGAGGCCCTAGGCGCAGAAATAGAGATTCCCTTGGACACACCTCCCTGTGTTCA 1865
DB 1501 AAGTAGAGAGGCCCTAGGCGCAGAAATAGAGATTCCCTTGGACACACCTCCCTGTGTTCA 1560
QY 1866 CTTTGTCTACAAGTAGGAGACACAGATGGACCTGTGGCCAGAGCAGCCTCAGACCTCC 1925
DB 1561 CTTTGTCTACAAGTAGGAGACACAGATGGACCTGTGGCCAGAGCAGCCTCAGACCTCC 1620
QY 1926 CCACCCACCANAATGCTCTGCTTGCCTTGATGAGAGAGAAAGGCTGGCAAGGTGGTTCAG 1985
DB 1621 CCACCCACCANAATGCTCTGCTTGCCTTGATGAGAGAGAAAGGCTGGCAAGGTGGTTCAG 1680
QY 1986 GGAATCTACTCTGTAGGAAACAGAAAAGAGAAAGAACGACTCTGCTGGCGGGAATACT 2045
DB 1681 GGAATCTACTCTGTAGGAAACAGAAAAGAGAAAGAACGACTCTGCTGGCGGGAATACT 1740
QY 2046 CTTGTGCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGTGAACTTCAGCCCTGAACCT 2105
DB 1741 CTTGTGCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGTGAACTTCAGCCCTGAACCT 1800
QY 2106 TTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGATTTCTTTTCTTAGTTTCAGAA 2165
DB 1801 TTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGATTTCTTTTCTTAGTTTCAGAA 1860
QY 2166 GTACTGGCATCACAGCAGTTTACCTTTGGCGTGTGTCCTGTGTCCTTGGTACCCCTGGCAGAGAAG 2225
DB 1861 GTACTGGCATCACAGCAGTTTACCTTTGGCGTGTGTCCTGTGTCCTTGGTACCCCTGGCAGAGAAG 1920
QY 2226 AGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAAGTAGGAGAGGATGACAGTTTGTCTATT 2285
DB 1921 AGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAAGTAGGAGAGGATGACAGTTTGTCTATT 1980
QY 2286 TGCCTTTAGACAGAGGACTGTATAAACAAGCCTTAACATTTGGTCAAGATTGCTCTTGA 2345
DB 1981 TGCCTTTAGACAGAGGACTGTATAAACAAGCCTTAACATTTGGTCAAGATTGCTCTTGA 2040
QY 2346 ATT 2348
DB 2041 ATT 2043
```

RESULT 10

ABL52457
ID ABL52457 standard; cDNA; 2070 BP.

AC ABL52457;

XX DT 16-JUL-2002 (first entry)

XX DE Human Asp-2(a) nucleotide sequence SEQ ID NO:3.

XX DE Human; Asp-1; Asp-2; aspartyl protease; enzyme; Alzheimer's disease;
KW proteolytic; chromosome 11q23.3-24.1; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT CDS 1..1506

FT FT /*tag= a
FT FT /product= "Asp-2(a)"
FT FT /note= "aspartyl protease"

XX GB2367060-A.

XX 27-MAR-2002.

PD

```
XX 29-OCT-2001; 2001GB-0025934.
XX 23-SEP-1999; 99US-155493P.
PR 23-SEP-1999; 99US-0404133.
PR 23-SEP-1999; 99WO-US20881.
PR 13-OCT-1999; 99US-0416901.
PR 06-DEC-1999; 99US-169232P.
PR 22-SEP-2000; 2000GB-0023315.
XX (PHRA ) PHARMACIA & UPJOHN CO.
XX PI Bienkowski MJ, Gurney M;
XX WPI; 2002-396337/43.
XX P-PSDB; ABB78590.
XX Human aspartyl protease 1 substrates useful in assays to detect
PT aspartyl protease activity, e.g. for the diagnosis of Alzheimer's
PT disease -
XX Example 2; Fig 2; 182pp; English.
XX The present invention describes a human aspartyl protease 1 (hu-Asp1)
CC substrate (I) which comprises a peptide of no more than 50 amino acids,
CC and which comprises the 8 amino acid sequence Gly-Leu-Ala-Leu-Ala-Leu-
CC Glu-Pro. Also described are: (1) a method (II) for assaying hu-Asp1
CC proteolytic activity, comprising: (a) contacting a hu-Asp1 protein with
CC (1) under acidic conditions; and (b) determining the level of hu-Asp1
CC proteolytic activity; (2) a purified polynucleotide (III) comprising a
CC nucleotide sequence that hybridises under stringent conditions to the
CC non-coding strand complementary to a defined 1804 nucleotide sequence
CC (see ABL52456) where the nucleotide sequence encodes a polypeptide having
CC Asp1 proteolytic activity and lacks nucleotides encoding a transmembrane
CC domain); (3) a purified polynucleotide (III') comprising a sequence that
CC hybridises under stringent conditions to (III) (the nucleotide sequence
CC encodes a polypeptide further lacking a pro-peptide domain corresponding
CC to amino acids 23-62 of hu-Asp1 (see ABB7859)); (4) a vector (IV)
CC comprising (III) or (III'); and (5) a host cell (V) transformed or
CC transfected with (III), (III') and/or (IV). The hu-Asp1 protease
CC substrate (I) may be used as an enzyme substrate in assays to detect
CC aspartyl protease activity, (II) and therefore diagnose diseases
CC associated with aberrant hu-Asp1 expression and activity such as
CC Alzheimer's disease. Hu-Asp1 has been localised to chromosome 21, while
CC hu-Asp2 has been localised to chromosome 11q23.3-24.1. The present
CC sequence encodes hu-Asp2(a) from the present invention.
XX Sequence 2070 BP; 476 A; 583 C; 562 G; 449 T; 0 other;
```

Query Match 86.9%; Score 2039.8; DB 24; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 306 ATGCCCAAGCCCTGCGCTGCTCTCTGATGGCGGGGAGTGTGCTGCCAC 365
DB 1 ATGCCCAAGCCCTGCGCTGCTCTCTGATGGCGGGGAGTGTGCTGCCAC 60

QY 366 GGCACCCAGCAGCATCCGCTGCCCTTGGCAGCGGCTGGGGGGCGCCCTGGG 425
DB 61 GGCACCCAGCAGCATCCGCTGCCCTTGGCAGCGGCTGGGGGGCGCCCTGGG 120

QY 426 CTGCGGCTGCCCGGGAGACCGAGAGCCCGAGGAGCCCGCCGAGGGGCGAGCTTT 485
DB 121 CTGCGGCTGCCCGGGAGACCGAGAGCCCGAGGAGCCCGCCGAGGGGCGAGCTTT 180

QY 486 GTGGAGATGTGGACAACCTGAGGGGCAAGTGGGGGAGGGCTACTACGTGGAGATGACC 545
DB 181 GTGGAGATGTGGACAACCTGAGGGGCAAGTGGGGGAGGGCTACTACGTGGAGATGACC 240

QY 546 GTGGGAGCCCCCGCAGAGCTCAACATCTGTGTGATACAGCAGCAGTAACCTTGA 605
DB 241 GTGGGAGCCCCCGCAGAGCTCAACATCTGTGTGATACAGCAGCAGTAACCTTGA 300

PA (PHAA) PHARMACIA & UPJOHN CO.
XX Yan R. Tomasselli AG, Gurney ME, Emmons TL, Bienkowski MJ;
PI Heinrikson RL;
XX WPI; 2002-216995/27.
DR P-PSDB; ABB06409.
XX Novel substrates for human aspartyl protease useful for identifying
PT modulators of beta secretase activity of aspartyl protease for treating
PT Alzheimer's disease
XX Claim 1; Page 117; 188pp; English.
XX The present invention describes an isolated peptide (I) comprising a
CC sequence of at least four amino acids, where the peptide is a substrate
CC for conducting aspartyl protease assays. (I) has neuroprotective and
CC nootropic activities, and can be used as an inhibitor of beta-secretase
CC activity. A beta-secretase modulator from the present invention can be
CC used for inhibiting beta-secretase activity in vivo, and in the
CC manufacture of a medicament for the treatment of Alzheimer's disease.
CC Pharmaceutical compositions from the present invention can be used for
CC treating a disease or condition characterised by an abnormal beta-
CC secretase activity. (I) is useful for identifying agents that modulate
CC the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful
CC as a core structure to construct derivatives. ABL49914 to ABL49925 and
CC ABB06409 to ABB06593 represent sequences used in the exemplification
CC of the present invention.
XX Sequence 2070 BP; 476 A; 583 C; 562 G; 449 T; 0 other;
Query Match 86.9%; Score 2039.8; DB 24; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 306 ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGGGGAGTGCCTGCCAC 365
Db 1 ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGGGGAGTGCCTGCCAC 60
Qy 366 GGCACCCAGCAGCATCCGCTGCCCTGGCGAGCGCTGGGGGGCCCCCTGGGG 425
Db 61 GGCACCCAGCAGCATCCGCTGCCCTGGCGAGCGCTGGGGGGCCCCCTGGGG 120
Qy 426 CTGGGGCTGCCCGGGACACGACGAGAGCCGAGGAGCCGCGGGAGGGGACGCTT 485
Db 121 CTGGGGCTGCCCGGGACACGACGAGAGCCGAGGAGCCGCGGGAGGGGACGCTT 180
Qy 486 GTGGAGATGGTGGACAACTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC 545
Db 181 GTGGAGATGGTGGACAACTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC 240
Qy 546 GTGGGACGCCCCGACAGCCTCAACATCCTGGTGATACAGGACGAGTAACCTTCCA 605
Db 241 GTGGGACGCCCCGACAGCCTCAACATCCTGGTGATACAGGACGAGTAACCTTCCA 300
Qy 606 GTGGGTGTCGCCGCCCCACCCCTTCCTGCGATCGCTACTACAGAGGACGCTGTCCAGCACA 665
Db 301 GTGGGTGTCGCCGCCCCACCCCTTCCTGCGATCGCTACTACAGAGGACGCTGTCCAGCACA 360
Qy 666 TACGGGACCTCGGGAAGGGTGTGTATGTGCCCTTACACCCAGGGCAATGGGAAGGGGAG 725
Db 361 TACGGGACCTCGGGAAGGGTGTGTATGTGCCCTTACACCCAGGGCAATGGGAAGGGGAG 420
Qy 726 CTGGGACCGACCTGGTAAAGTATCCCCATGCCCCACAGCTCAGTGTCCGTCACACATT 785
Db 421 CTGGGACCGACCTGGTAAAGTATCCCCATGCCCCACAGCTCAGTGTCCGTCACACATT 480
Qy 786 GCTGCATCACTGAATCAGAAAGTTCTTCATCAACGGCTCCAAGTGGGAGGCAATCCTG 845
Db 481 GCTGCATCACTGAATCAGAAAGTTCTTCATCAACGGCTCCAAGTGGGAGGCAATCCTG 540
Qy 846 GGGCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTCTTTGACTCT 905
Db 1621 CCACCCACCAATGCTCTGCTGCTTGTATGAGAAAGGCTGGCAAGGCTGGGTTCCAG 1680

Db 541 GGGCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTCTTTGACTCT 600
Qy 906 CTGGTAAAGCAGACCCACGTTCCCAACCTTCTTCCCTGCGAGCTTTGGTCTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACGTTCCCAACCTTCTTCCCTGCGACCTTTGGTCTGGCTTC 660
Qy 966 CCCCTCAACCACTGTGAAGTGTGGCCTCTGTGCGAGGAGGAGCATGATCATTTGGAGTATC 1025
Db 661 CCCCTCAACCACTGTGAAGTGTGGCCTCTGTGCGAGGAGGAGCATGATCATTTGGAGTATC 720
Qy 1026 GACCACTCGCTGTACACAGGAGTCTCTGGTATACACCATCCCGGGGAGTGGTATAT 1085
Db 721 GACCACTCGCTGTACACAGGAGTCTCTGGTATACACCATCCCGGGGAGTGGTATAT 780
Qy 1086 GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGACATGCAAGGAG 1145
Db 781 GAGGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAAATGACATGCAAGGAG 840
Qy 1146 TACAACTATGACAAGAGCATTTGTGGACAGTGGCACACCAACCTTCGTTTGGCCCAAGAA 1205
Db 841 TACAACTATGACAAGAGCATTTGTGGACAGTGGCACACCAACCTTCGTTTGGCCCAAGAA 900
Qy 1206 GTGTTTGAAGTGCAGTCAATCCATCAAGGAGGCTCTCCACGAGAGTTCCTGTAT 1265
Db 901 GTGTTTGAAGTGCAGTCAATCCATCAAGGAGGCTCTCCACGAGAGTTCCTGTAT 960
Qy 1266 GGTTCCTGCTAGGAGAGCAGCTGCTGTGTCGCAAGCAGGACCAACCTTCGGAACATT 1325
Db 961 GGTTCCTGCTAGGAGAGCAGCTGCTGTGTCGCAAGCAGGACCAACCTTCGGAACATT 1020
Qy 1326 TTCCAGTCACTCTACTTACCTAATGGGTGAGGTTACCAACAGTTCCTTCGCGATCAC 1385
Db 1021 TTCCAGTCACTCTACTTACCTAATGGGTGAGGTTACCAACAGTTCCTTCGCGATCAC 1080
Qy 1386 ATCTTCGCGAGCAATACCTGCGGCCAGTGGGAAGATGGGCCACCTCCCAAGACACTGT 1445
Db 1081 ATCTTCGCGAGCAATACCTGCGGCCAGTGGGAAGATGGGCCACCTCCCAAGACACTGT 1140
Qy 1446 TACAACTTGGCATCTACAGTCATCCAGGACAGTGTATGGAGCTGTATCATGGAG 1505
Db 1141 TACAACTTGGCATCTACAGTCATCCAGGACAGTGTATGGAGCTGTATCATGGAG 1200
Qy 1506 GGTTCCTACGTTGTCTTTGATCGGGCCGAAACAAATTTGGCTTCTGCTCAGCGCTTCG 1565
Db 1201 GGTTCCTACGTTGTCTTTGATCGGGCCGAAACAAATTTGGCTTCTGCTCAGCGCTTCG 1260
Qy 1566 CATGTGCAGATGAGTTTCAGGACGCGAGCGGTGGAGGCGCTTTTGTACCTTTGGACATG 1625
Db 1261 CATGTGCAGATGAGTTTCAGGACGCGAGCGGTGGAGGCGCTTTTGTACCTTTGGACATG 1320
Qy 1626 GAAGACTGTGGCTACAACATTCACAGACAGATGATCAACCTCATGACCATAGCCTAT 1685
Db 1321 GAAGACTGTGGCTACAACATTCACAGACAGATGATCAACCTCATGACCATAGCCTAT 1380
Qy 1686 GTCATGGTGGCATCTGCGCCCTCTTCATGTGCCACTCTGCTCATGTGTGTGTCAGTGG 1745
Db 1381 GTCATGGTGGCATCTGCGCCCTCTTCATGTGCCACTCTGCTCATGTGTGTGTCAGTGG 1440
Qy 1746 CGCTCCCTCCGCTGCTGCGCCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTCCCTCCGCTGCTGCGCCAGCAGCATGATGACTTTTGTGTGATGACATCTCCCTGCTG 1500
Qy 1806 AAGTCAGAGGCGCCATGGCAGAAAGATAGAGATTCCTGGACCAACCTCCGTGGTTCA 1865
Db 1501 AAGTCAGAGGCGCCATGGCAGAAAGATAGAGATTCCTGGACCAACCTCCGTGGTTCA 1560
Qy 1866 CTTTGGTCAACAGTAGGAGACACAGATGCGACCTGTGGCGAGAGCACCTTCAGGACCTCC 1925
Db 1561 CTTTGGTCAACAGTAGGAGACACAGATGCGACCTGTGGCGAGAGCACCTTCAGGACCTCC 1620
Qy 1926 CCACCCACCAATGCTCTGCTGCTTGTATGAGAGGAAAGGCTGGCAAGGCTGGGTTCCAG 1985
Db 1621 CCACCCACCAATGCTCTGCTGCTTGTATGAGAGGAAAGGCTGGCAAGGCTGGGTTCCAG 1680

Qy 1206 GTGTTTGAAGCTGCAGTCAATCATCAAGCAGCCTCCTCCACGGAGAAGTTCCCTGAT 1265
Db 901 GTGTTTGAAGCTGCAGTCAATCATCAAGCAGCCTCCTCCACGGAGAAGTTCCCTGAT 960
Qy 1266 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGTGGCAAGCAGCACCCTTGGAAACATT 1325
Db 961 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGTGGCAAGCAGCACCCTTGGAAACATT 1020
Qy 1326 TTCCCACTCATCTACCTACCTAATGGGTGAGGTACCAACACAGTCTTCCCGATCACC 1385
Db 1021 TTCCCACTCATCTACCTACCTAATGGGTGAGGTACCAACACAGTCTTCCCGATCACC 1080
Qy 1386 ATCTCTCCGAGCAATACCTCGGCGCAGTGAAGATGTGGCCACGTCCCAAGACGACTGT 1445
Db 1081 ATCTCTCCGAGCAATACCTCGGCGCAGTGAAGATGTGGCCACGTCCCAAGACGACTGT 1140
Qy 1446 TACAAGTTTGGCATCTCACAGTCATCACAGGGGCACTGTATGGAGAGTGTATCATGGAG 1505
Db 1141 TACAAGTTTGGCATCTCACAGTCATCACAGGGGCACTGTATGGAGAGTGTATCATGGAG 1200
Qy 1506 GGCTTCTACGTTGCTTTGTATCGGGGCGGAAACGAATTGGCTTTGCTGTCAGGCGCTGC 1565
Db 1201 GGCTTCTACGTTGCTTTGTATCGGGGCGGAAACGAATTGGCTTTGCTGTCAGGCGCTGC 1260
Qy 1566 CATGTGCACGATGATTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACACCTTGGACATG 1625
Db 1261 CATGTGCACGATGATTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACACCTTGGACATG 1320
Qy 1626 GAAGACTGTGGCTACAACTATCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
Db 1321 GAAGACTGTGGCTACAACTATCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380
Qy 1686 GTCATGCTGCCATCTCGGCCCTTTCATGCTGCCATCTGCCATCATGGTGTGTCAGTGG 1745
Db 1381 GTCATGCTGCCATCTCGGCCCTTTCATGCTGCCATCTGCCATCATGGTGTGTCAGTGG 1440
Qy 1746 CGCTGCCCTCCGCTCGCTCGCCACGACGATGATGACTTTGCTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGCCCTCCGCTCGCTCGCCACGACGATGATGACTTTGCTGATGACATCTCCCTGCTG 1500
Qy 1806 AAGTGAGAGGCCCATGGGCAAGAGATAGAGATTCCTCCCT -GGACACACCTCCGTTGTTTC 1864
Db 1501 AAGTGAGAGGCCCATGGGCAAGAGATAGAGATTCCTCCCTGGGAGCCACACCTCCGTTGTTTC 1560
Qy 1865 ACTTTGCTCAAGTACGAGACAGATGGCACCCTGTGGCCAGAGCACCCTCAGGACCCCTC 1924
Db 1561 ACTTTGCTCAAGTACGAGACAGATGGCACCCTGTGGCCAGAGCACCCTCAGGACCCCTC 1620
Qy 1925 CCCACCACCAATTCCTCTGCTTGATGGAGAGAAAGGCTGGCAAGGTGGGTTCCA 1984
Db 1621 CCCACCACCAATTCCTCTGCTTGATGGAGAGAAAGGCTGGCAAGGTGGGTTCCA 1680
Qy 1985 GGGACTGTACCTGTAGGAACAGAAAAGAGAAAGACACATCTGCTGGCGGGAATAC 2044
Db 1681 GGGACTGTACCTGTAGGAACAGAAAAGAGAAAGACACATCTGCTGGCGGGAATAC 1740
Qy 2045 TCTTGGTCACTCAAAATTAAGTCGGGAAATTCCTGCTTGAACATTCAGCCCTGAACC 2104
Db 1741 TCTTGGTCACTCAAAATTAAGTCGGGAAATTCCTGCTTGAACATTCAGCCCTGAACC 1800
Qy 2105 TTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATTCTTCTTTTAGTTTTCAGA 2164
Db 1801 TTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATTCTTCTTTTAGTTTTCAGA 1860
Qy 2165 AGTACTGGCATCACACGAGGTTTACCTTGGCGTGTGTCCTGTGGTACCCCTGGCAGAGAA 2224
Db 1861 AGTACTGGCATCACACGAGGTTTACCTTGGCGTGTGTCCTGTGGTACCCCGGCGAGAGAA 1920
Qy 2225 GAGACCAAGCTTGTTCCTGCTGGCCAAAGTTCAGTAGGAGGATGCACAGTTTGTAT 2284
Db 1921 GAGACCAAGCTTGTTCCTGCTGGCCAAAGTTCAGTAGGAGGATGCACAGTTTGTAT 1980

Qy 2285 TTGCTTTAGACACAGGACTGTATATAACAGCCTTAACATTGGTGCAAGATTGGCTCTTG 2344
Db 1981 TTGCTTTAGACACAGGACTGTATATAACAGCCTTAACATTGGTGCAAGATTGGCTCTTG 2040
Qy 2345 AATT 2348
Db 2041 AATT 2044
RESULT 13
AAF31848
ID AAF31848 standard; cDNA; 3252 BP.
XX AAF31848;
XX AC
XX DT 12-APR-2001 (first entry)
XX Human memapsin 2 cDNA.
XX Human; memapsin 2; nootropic; neuroprotective; amyloid precursor protein;
KW APP; memapsin 2 inhibitor; Alzheimer's disease; ds.
XX Homo sapiens.
OS WO200100665-A2.
PN 04-JAN-2001.
PD 27-JUN-2000; 2000WO-US17742.
XX 28-JUN-1999; 99US-0141363.
PR 30-NOV-1999; 99US-0168060.
PR 25-JAN-2000; 2000US-0177836.
PR 27-JAN-2000; 2000US-0178368.
PR 08-JUN-2000; 2000US-0210292.
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
PA (UNII) UNIV ILLINOIS FOUND.
XX Tang JJN, Hong L, Ghosh AK;
XX WPI; 2001-137933/14.
DR P-PSDB; AAB66572.
XX Novel memapsin 2 inhibitors which bind to active site of memapsin 2
PT having 2 catalytic aspartic residues and substrate binding cleft, used
PT to treat Alzheimer's disease by blocking amyloid precursor protein
PT cleavage
XX
PS Example 1; Page 70-71; 86pp; English.
XX The present sequence is given in a specification relating to an inhibitor
CC of catalytically active memapsin 2. The inhibitor binds to the memapsin 2
CC active site, which is defined by the presence of two catalytic aspartic
CC residues and a substrate binding cleft. The inhibitor is useful for
CC the treatment and diagnosis of Alzheimer's disease. It is useful in
CC screens for individuals with a genetic predisposition to Alzheimer's
CC disease. The inhibitor is useful as a reagent for specifically binding to
CC memapsin 2 or memapsin 2 analogues and for aiding in memapsin 2
CC isolation, purification and characterisation.
XX
SQ Sequence 3252 BP; 804 A; 863 C; 811 G; 771 T; 3 other;
Query Match 85.3%; Score 2004; DB 22; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 345 GCGGAGTGCTGCTGCCACCGGACCCAGCAGCGCATCCGGCTGCCCTCGCAGCGGC 404
Db 1 GCGGAGTGCTGCTGCCACCGGACCCAGCAGCGCATCCGGCTGCCCTCGCAGCGGC 60
Qy 405 CTGGGGGCGCCCCCTCGGGCTCGGGCTGCCCGGGAGACCGAGAGCCCGAGGAG 464
|||||

XX 04-JAN-2001.
XX 27-JUN-2000; 2000WO-US17661.
XX
XX 28-JUN-1999; 99US-0141363.
PR 30-NOV-1999; 99US-0168060.
PR 25-JAN-2000; 2000US-0177836.
PR 27-JAN-2000; 2000US-0178368.
PR 08-JUN-2000; 2000US-0210292.
XX
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
XX Tang JJN, Lin X, Koelsch G;
PI WPI; 2001-102885/11.
DR
XX Purified recombinant catalytically active memapsin 2, used to screen
PT inhibitors of it, which are used to treat and prevent Alzheimer's
PT disease -
XX
XX Example 1; Page 71-72; 86pp; English.
XX
CC The present invention relates to a purified recombinant
CC catalytically active memapsin 2. The invention may be used for
CC isolating inhibitors which are used to treat or prevent
CC Alzheimer's disease. The invention may also be used to screen
CC for individuals more genetically prone to develop Alzheimer's
CC disease.
XX
XX Sequence 3252 BP; 804 A; 863 C; 811 G; 771 T; 3 other;
SQ

Query Match 85.3%; Score 2004; DB 22; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 GCGGGAGTGTGCTGCCCGCCAGCCAGCAGCGCATCCGGCTGCCCTGCCAGCGGC 404
DB 1 GCGGGAGTGTGCTGCCCGCCAGCCAGCAGCGCATCCGGCTGCCCTGCCAGCGGC 60
QY 405 CTGGGGGGCGCCCTTGGGGCTGGCGCTGCCCGGGAGACCGAGAGCCCGAGGAG 464
DB 61 CTGGGGGGCGCCCTTGGGGCTGGCGCTGCCCGGGAGACCGAGAGCCCGAGGAG 120
QY 465 CCCGGCCGAGGGGACGCTTGTGGAGATGGTGACACCTGAGGGGCAAGTCGGGGCAG 524
DB 121 CCCGGCCGAGGGGACGCTTGTGGAGATGGTGACACCTGAGGGGCAAGTCGGGGCAG 180
QY 525 GGCTACTAGTGGAGATGACCGTGGGGAGCCCGCCCGAGAGCGCTCAACATCCTGGTGGAT 584
DB 181 GGCTACTAGTGGAGATGACCGTGGGGAGCCCGCCCGAGAGCGCTCAACATCCTGGTGGAT 240
QY 585 ACAGGAGCAGTAACTTTGAGTGGGTGGCTGCCCGCCCGCCCGCCCTTCCCTGCATCGTACTAC 644
DB 241 ACAGGAGCAGTAACTTTGAGTGGGTGGCTGCCCGCCCGCCCGCCCTTCCCTGCATCGTACTAC 300
QY 645 CAGAGGAGCTGTCCAGCACATACCGGACCTCCGGAGGGTGTGTATGTGCCCTACACC 704
DB 301 CAGAGGAGCTGTCCAGCACATACCGGACCTCCGGAGGGTGTGTATGTGCCCTACACC 360
QY 705 CAGGGCAAGTGGGAAGGGAGCTGGGACCGACCTGTGTAGCATCCCGCCATGGCCCCCAAC 764
DB 361 CAGGGCAAGTGGGAAGGGAGCTGGGACCGACCTGTGTAGCATCCCGCCATGGCCCCCAAC 420
QY 765 GTCACTGTGGTGGCAACATTTGCCATCTCACTGAATCAGACAAGTTCTTCATCAACGGC 824
DB 421 GTCACTGTGGTGGCAACATTTGCCATCTCACTGAATCAGACAAGTTCTTCATCAACGGC 480
QY 825 TCCAACCTGGGAAGGCATCTTGGGGCTGGCCCTATGCTGAGATTGCCAGGCTGAGACTCC 884
DB 481 TCCAACCTGGGAAGGCATCTTGGGGCTGGCCCTATGCTGAGATTGCCAGGCTGAGACTCC 540
QY 885 CTGAGGCTTTCTTTGACTCTCTGGTAAAGCAGACCCACCGTTCCCAACCTCTCTCCCTG 944

DB 541 CTGAGGCTTTCTTTGACTCTCTGTTAAAGCAGACCCACGTTCCCAACCTCTCTCCCTG 600
QY 945 CAGCTTTGTGGTGTGGCTTCCCTCAACAGCTCTGAAGTGTGGCTCTGTGGAGGG 1004
DB 601 CAGCTTTGTGGTGTGGCTTCCCTCAACAGCTCTGAAGTGTGGCTCTGTGGAGGG 660
QY 1005 AGCATGATCATTGGAGGTATCGAGCACTCGCTGTACAGCAGCTCTCTGGTATACACCC 1064
DB 661 AGCATGATCATTGGAGGTATCGAGCACTCGCTGTACAGCAGCTCTCTGGTATACACCC 720
QY 1065 ATCCGGCGGAGTGGTATTATGAGGTGATCATTTGGGGTGGAGATCAATGGACAGAT 1124
DB 721 ATCCGGCGGAGTGGTATTATGAGGTGATCATTTGGGGTGGAGATCAATGGACAGAT 780
QY 1125 CTGAAATGGAAGTCAAGGAGTCAACTATGACAGAGCAATTTGGAGCTGGCAGCACC 1184
DB 781 CTGAAATGGAAGTCAAGGAGTCAACTATGACAGAGCAATTTGGAGCTGGCAGCACC 840
QY 1185 AACCTTGGTTGCCCAAGAAAGTGTGGAAGCTGAGTCAAAATCCATCAAGGACGCTCC 1244
DB 841 AACCTTGGTTGCCCAAGAAAGTGTGGAAGCTGAGTCAAAATCCATCAAGGACGCTCC 900
QY 1245 TCCACGAGAAAGTCCCTGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGTGCAAGCA 1304
DB 901 TCCACGAGAAAGTCCCTGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGTGCAAGCA 960
QY 1305 GGCACACCCCTTGGAAACATTTCCAGTCACTCTACCTTAATGGGTGAGTTACC 1364
DB 961 GGCACACCCCTTGGAAACATTTCCAGTCACTCTACCTTAATGGGTGAGTTACC 1020
QY 1365 AACAGTCCCTTCCGATCACCATCTCCGAGCAATACCTCGGCGCAGTGAAGATGTG 1424
DB 1021 AACAGTCCCTTCCGATCACCATCTCCGAGCAATACCTCGGCGCAGTGAAGATGTG 1080
QY 1425 GCCACGTCCTTCCAGAGCAGCTGTTTACAAGTTTCCAGTCTCACAGTCACTCACGGCAGCTGT 1484
DB 1081 GCCACGTCCTTCCAGAGCAGCTGTTTACAAGTTTCCAGTCTCACAGTCACTCACGGCAGCTGT 1140
QY 1485 ATGGAGCTGTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCGAAGCAAT 1544
DB 1141 ATGGAGCTGTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCGAAGCAAT 1200
QY 1545 GGCCTTGTGTGACGGCTTGCATGTGCAGCATGAGTTTCAGGAGCGGCGTGAAGGC 1604
DB 1201 GGCCTTGTGTGACGGCTTGCATGTGCAGCATGAGTTTCAGGAGCGGCGTGAAGGC 1260
QY 1605 CCTTTGTGACCTTGGACATGGAAGCTGTGGCTACAACTTCCACAGACAGATGAGTCA 1664
DB 1261 CCTTTGTGACCTTGGACATGGAAGCTGTGGCTACAACTTCCACAGACAGATGAGTCA 1320
QY 1665 ACCCTCATGACCATAGCTATGTCATGGCTGCCATCTCGCCCTCTTCTATGCTGCCACTC 1724
DB 1321 ACCCTCATGACCATAGCTATGTCATGGCTGCCATCTCGCCCTCTTCTATGCTGCCACTC 1380
QY 1725 TGCCTCATGCTGTGTGAGTGGCGTGCCTCCGCTGCCCTGCCAGCAGCATGATGACTTT 1784
DB 1381 TGCCTCATGCTGTGTGAGTGGCGTGCCTCCGCTGCCCTGCCAGCAGCATGATGACTTT 1440
QY 1785 GCTGATGACATCTCCCTGCTGAAAGTGGAGGGCCCATGGGAGAGATAGAGATTCCTCT 1844
DB 1441 GCTGATGACATCTCCCTGCTGAAAGTGGAGGGCCCATGGGAGAGATAGAGATTCCTCT 1500
QY 1845 GGACACACCTCCGTTGGTTCATTTGGTTCACAAGTAGGAGACACAGATGGCACCTGTGGC 1904
DB 1501 GGACACACCTCCGTTGGTTCATTTGGTTCACAAGTAGGAGACACAGATGGCACCTGTGGC 1560
QY 1905 CAGAGCAGCTCAGGACCTCCCGCCACCAATGCCCTCTGCTTGTGAGGAGAGAGAAA 1964
DB 1561 CAGAGCAGCTCAGGACCTCCCGCCACCAATGCCCTCTGCTTGTGAGGAGAGAGAAA 1620
QY 1965 GGCTGGCAAGTGGGTTCAGGGGACTGTACCTGTAGGAAACAGAGAGAGAGAAAG 2024

QY 1125 CTGAAATGAGCTGCAAGGAGTACAACTATGACAAGAGCATTTGGACAGTGGCACCACC 1184
DB 781 CTGAAATGAGCTGCAAGGAGTACAACTATGACAAGAGCATTTGGACAGTGGCACCACC 840
QY 1185 AACCTTCGTTTGGCCCAAGAAAGTGTGTAAGCTGCAAGTCAAAATCCATCAAGGAGGCTCC 1244
DB 841 AACCTTCGTTTGGCCCAAGAAAGTGTGTAAGCTGCAAGTCAAAATCCATCAAGGAGGCTCC 900
QY 1245 TCCACGAGGAAGTTCCTTGATGGTTTCTGCTAGGAGAGCAGCTGTGTGCTGGCAAGCA 1304
DB 901 TCCACGAGGAAGTTCCTTGATGGTTTCTGCTAGGAGAGCAGCTGTGTGCTGGCAAGCA 960
QY 1305 GGCACACCCCTTGGAAACATTTTCCAGTCACTCTACTTACCTAATGGGTGAGGTTACC 1364
DB 961 GGCACACCCCTTGGAAACATTTTCCAGTCACTCTACTTACCTAATGGGTGAGGTTACC 1020
QY 1365 AACCACTCTTCCGGATCACCATCCTTCCGACGCAATACCTGGCGCCAGTGGGAAGATGTG 1424
DB 1021 AACCACTCTTCCGGATCACCATCCTTCCGACGCAATACCTGGCGCCAGTGGGAAGATGTG 1080
QY 1425 GCCACGTCCCAAGCAGCTGTTTACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTT 1484
DB 1081 GCCACGTCCCAAGCAGCTGTTTACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTT 1140
QY 1485 ATGGAGCTGTTATCATGGAGGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAAAT 1544
DB 1141 ATGGAGCTGTTATCATGGAGGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAAAT 1200
QY 1545 GGCTTTGCTGTACGCGCTTGCATGTGCAGGATGAGTTTCAGGACGGCAGCGGTGGGAAGGC 1604
DB 1201 GGCTTTGCTGTACGCGCTTGCATGTGCAGGATGAGTTTCAGGACGGCAGCGGTGGGAAGGC 1260
QY 1605 CCTTTTGTCACTTGGACATGGAAGACTTGGCTACAACATCCACAGACAGATGAGTCA 1664
DB 1261 CCTTTTGTCACTTGGACATGGAAGACTTGGCTACAACATCCACAGACAGATGAGTCA 1320
QY 1665 ACCCTCATGACCATAGCCTATGTATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTC 1724
DB 1321 ACCCTCATGACCATAGCCTATGTATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTC 1380
QY 1725 TGGCTCATGGTGTGTCAGTGGCGCTGCCCTCCGCTGCCCTGGCCAGCAGCATGACTTT 1784
DB 1381 TGGCTCATGGTGTGTCAGTGGCGCTGCCCTCCGCTGCCCTGGCCAGCAGCATGACTTT 1440
QY 1785 GCTGATGACATCTCCCTGCTGTAAGTGGAGGCCCCATGGGCAAGATAGAGATTCCOCT 1844
DB 1441 GCTGATGACATCTCCCTGCTGTAAGTGGAGGCCCCATGGGCAAGATAGAGATTCCOCT 1500
QY 1845 GGACCACACCTCCGTTGGTTTCACTTTGGTCAAAAGTAGGAGACACAGATGCCACTGTGGC 1904
DB 1501 GGACCACACCTCCGTTGGTTTCACTTTGGTCAAAAGTAGGAGACACAGATGCCACTGTGGC 1560
QY 1905 CAGAGCACCTCAGGACCTCCGCCACCACCAATGGCTGTGCCCTTGTAGGAGAGGAAAA 1964
DB 1561 CAGAGCACCTCAGGACCTCCGCCACCACCAATGGCTGTGCCCTTGTAGGAGAGGAAAA 1620
QY 1965 GGCTGGCAGAGTGGGTTCCAGGACTGTACTCTAGSAAACAGAAAGAGAAAGAAAG 2024
DB 1621 GGCTGGCAGAGTGGGTTCCAGGACTGTACTCTAGSAAACAGAAAGAGAAAGAAAG 1680
QY 2025 CACTCTGCTGGCGGGAATACTCTTGTCTACCTCAAAATTTAAAGTCGGGAAATCTGCTGCT 2084
DB 1681 CACTCTGCTGGCGGGAATACTCTTGTCTACCTCAAAATTTAAAGTCGGGAAATCTGCTGCT 1740
QY 2085 TGAACCTTCAGCCCTGAACCTTTGTCCACCAATTCCTTTAAATTTCCAAACCCAAAGTATT 2144
DB 1741 TGAACCTTCAGCCCTGAACCTTTGTCCACCAATTCCTTTAAATTTCCAAACCCAAAGTATT 1800
QY 2145 CTTCTTTTCTTAGTTTCAGAAAGTACTGGCATCACAGCAGGTTACCTTGGCGTGTGCC 2204
DB 1801 CTTCTTTTCTTAGTTTCAGAAAGTACTGGCATCACAGCAGGTTACCTTGGCGTGTGCC 1860
QY 2205 TGTGGTACCCTTGGCAGAGAAGACCAAGCTTTGTTCCCTGTGTGGCCAAAGTCAGTAGGA 2264

DB 1861 TGTGGTACCCTGGCAGAGAAGACCAAGCTTGTTCCTTGTGGCCAAAGTCAGTAGGA 1920
QY 2265 GAGGATGCACAGTTTGTCTATTTTGTCTTTAGAGACAGGAGCTGTATAACAAGCCTTAACATT 2324
DB 1921 GAGGATGCACAGTTTGTCTATTTTGTCTTTAGAGACAGGAGCTGTATAACAAGCCTTAACATT 1980
QY 2325 GGTGCAAAAGATTGCCCTCTTGAATT 2348
DB 1981 GGTGCAAAAGATTGCCCTCTTGAATT 2004

Search completed: March 1, 2003, 22:17:53
Job time : 353.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2003, 14:54:35 ; Search time 56 Seconds
(without alignments)
12858.519 Million cell updates/sec

Title: US-09-723-722A-44

Perfect score: 2348

Sequence: 1 ccatgcgcgcctcacagc.....caagattgccttgaatt 2348

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	2041.4	86.9	2070	4 US-09-548-372D-3
2	2041.4	86.9	2070	4 US-09-548-367D-3
3	2024	86.2	2541	4 US-09-009-191-1
4	1881	80.1	1977	4 US-09-548-372D-5
5	1881	80.1	1977	4 US-09-548-367D-5
6	1835.4	78.2	2370	4 US-09-009-191-3
7	1550	66.0	2043	4 US-09-548-372D-7
8	1550	66.0	2043	4 US-09-548-367D-7
9	1355.8	57.7	1362	4 US-09-548-372D-29
10	1355.8	57.7	1362	4 US-09-548-367D-29
11	1355.8	57.7	1380	4 US-09-548-372D-31
12	1355.8	57.7	1380	4 US-09-548-367D-31
13	1298	55.3	1506	4 US-09-713-158-1
14	1293.8	55.1	1341	4 US-09-548-372D-21
15	1293.8	55.1	1341	4 US-09-548-367D-21
16	1293.8	55.1	1380	4 US-09-548-372D-23
17	1293.8	55.1	1380	4 US-09-548-367D-23
18	1270.4	54.1	1302	4 US-09-548-372D-25
19	1270.4	54.1	1302	4 US-09-548-367D-25
20	1197.4	51.0	1287	4 US-09-548-372D-50
21	1197.4	51.0	1287	4 US-09-548-367D-50
22	1197.4	51.0	1305	4 US-09-548-372D-52
23	1197.4	51.0	1305	4 US-09-548-367D-52
24	1180	50.3	1278	4 US-09-548-372D-27
25	1180	50.3	1278	4 US-09-548-367D-27
26	421.2	17.9	511	4 US-09-280-116-30
27	397.4	16.9	1804	4 US-09-548-372D-1

ALIGNMENTS

RESULT 1

US-09-548-372D-3

: Sequence 3, Application US/09548372D

: Patent No. 6420534

: GENERAL INFORMATION:

: APPLICANT: GURNEY ET AL.

: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE

: FILE REFERENCE: 29915/62801

: CURRENT APPLICATION NUMBER: US/09/548,372D

: CURRENT FILING DATE: 2000-04-12

: PRIOR APPLICATION NUMBER: US 60/155,493

: PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: US 09/404,133

: PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: PCT/US99/20881

: PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: US 60/101,594

: PRIOR FILING DATE: 1998-09-24

: NUMBER OF SEQ ID NOS: 73

: SOFTWARE: PatentIn version 3.1

: SEQ ID NO 3

: LENGTH: 2070

: TYPE: DNA

: ORGANISM: Homo sapiens

US-09-548-372D-3

Query Match 86.9%; Score 2041.4; DB 4; Length 2070;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2042; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 306 ATGGCCCAAGCCCTGCCCTGGCTCCTGTGTGGATGGCGGGAGTGTGCTGCTGCCAC 365

Db 1 ATGGCCCAAGCCCTGCCCTGGCTCCTGTGTGGATGGCGGGAGTGTGCTGCTGCCAC 60

Qy 366 GGCACCCAGCAGCGGATCGGCTGCCCTGCCACGCGGCTGGGGGGCCCCCTGGGG 425

Db 61 GGCACCCAGCAGCGGATCGGCTGCCCTGCCACGCGGCTGGGGGGCCCCCTGGGG 120

Qy 426 CTGCGGCTGCCCGGGAGACCGACGAGAGCCCGGAGGAGCCCGGAGGCGAGCTTT 485

Db 121 CTGCGGCTGCCCGGGAGACCGACGAGAGCCCGGAGGAGCCCGGAGGCGAGCTTT 180

Qy 486 CTGGAGATGGTGGCAACCTAGGGGCAAGTGGGGCAGGCTACTTACCTGGAGATGACC 545

Db 181 GTGGAGATGGTGGCAACCTAGGGGCAAGTGGGGCAGGCTACTTACCTGGAGATGACC 240

Qy 546 GTGGCAGCCCCCGCAGCAGCTCAACATCTGTTGGATACAGGAGCAGTACCTTTGCA 605

Db 241 GTGGCAGCCCCCGCAGCAGCTCAACATCTGTTGGATACAGGAGCAGTACCTTTGCA 300

i: ORGANISM: Homo sapiens									
US-09-548-367D-3									
Query Match 86.9%; Score 2041.4; DB 4; Length 2070;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2042; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	306	ATGGCCCAAGCCCTGGCCCTGGCTCCTGCTGTGGATGGCGGGAGTGCTGCCCTGCCAC	365						
DB	1	ATGGCCCAAGCCCTGGCCCTGGCTCCTGCTGTGGATGGCGGGAGTGCTGCCCTGCCAC	60						
QY	366	GGCACCCAGCACGGCATCCGGCTGCCCTGGCAGCGCCCTGGGGGGGGCCCCCTGGGG	425						
DB	61	GGCACCCAGCACGGCATCCGGCTGCCCTGGCAGCGCCCTGGGGGGGGCCCCCTGGGG	120						
QY	426	CTGGGGTGGCCCCGGGAGACCGAAGAGCCCGAGGACCGCCGGCGAGGGGAGCTTT	485						
DB	121	CTGGGGTGGCCCCGGGAGACCGAAGAGCCCGAGGACCGCCGGCGAGGGGAGCTTT	180						
QY	486	GTGGAGATGTGTGAACACCTGAGGGGCAAGTCGGGGCAGGCTACTAGCTGGAGATGACC	545						
DB	181	GTGGAGATGTGTGAACACCTGAGGGGCAAGTCGGGGCAGGCTACTAGCTGGAGATGACC	240						
QY	546	GTGGGACGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAATTTGCA	605						
DB	241	GTGGGACGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAATTTGCA	300						
QY	606	GTGGGTCTGCCCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGCGACGTGTCAGACA	665						
DB	301	GTGGGTCTGCCCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGCGACGTGTCAGACA	360						
QY	666	TACCGGACCTCCGGAAAGGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGAG	725						
DB	361	TACCGGACCTCCGGAAAGGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGAG	420						
QY	726	CTGGGCACGACCTGGTAAGCATCCCCCATGGCCCCCAACGTCATGTGCGTGCACAATTT	785						
DB	421	CTGGGCACGACCTGGTAAGCATCCCCCATGGCCCCCAACGTCATGTGCGTGCACAATTT	480						
QY	786	GCTGCCATCACTGAATCAGACAAGTTCTTCAATCAAGGCTCCCACTGGGAAGGCATCTG	845						
DB	481	GCTGCCATCACTGAATCAGACAAGTTCTTCAATCAAGGCTCCCACTGGGAAGGCATCTG	540						
QY	846	GGCGTGCCCTATGCTGAGATTGCCAGGCTTGACGACCTCCCTGGAGCCCTTTCTTTGACTCT	905						
DB	541	GGCGTGCCCTATGCTGAGATTGCCAGGCTTGACGACCTCCCTGGAGCCCTTTCTTTGACTCT	600						
QY	906	CTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCCCTGCAGCTTTGTGTGCTGGCTTC	965						
DB	601	CTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCCCTGCAGCTTTGTGTGCTGGCTTC	660						
QY	966	CCCTCAACCACTGTAAGTGTGCGCTCTCTCGAGGGAGCATGATCATTTGGAGGTATC	1025						
DB	661	CCCTCAACCACTGTAAGTGTGCGCTCTCTCGAGGGAGCATGATCATTTGGAGGTATC	720						
QY	1026	GACCACTCGCTGTACACAGGCAGTCTCTGTGTATACACCCATCCGGCGGGAGTGTATTAT	1085						
DB	721	GACCACTCGCTGTACACAGGCAGTCTCTGTGTATACACCCATCCGGCGGGAGTGTATTAT	780						
QY	1086	GAGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	1145						
DB	781	GAGTGATCATTTGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG	840						
QY	1146	TACAACATATGACAGACATTTGACAGTGGCAGCCACCAACCTTCGTTTCCCAAGAAA	1205						
DB	841	TACAACATATGACAGACATTTGACAGTGGCAGCCACCAACCTTCGTTTCCCAAGAAA	900						
QY	1206	GTGTTTGAAGCTGCAGTCAATTCATCAAGCAGCCTCCTCCACGGGAAAGTTCCCTGAT	1265						
DB	901	GTGTTTGAAGCTGCAGTCAATTCATCAAGCAGCCTCCTCCACGGGAAAGTTCCCTGAT	960						
QY	1266	GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGCACCCCTTTGSAACATTT	1325						

RESULT 3

US-09-009-191-1

: Sequence 1, Application US/09009191

: Patent No. 6319689

: GENERAL INFORMATION:

: APPLICANT: POWELL, DAVID

: APPLICANT: CHAPMAN, CONRAD

: APPLICANT: MURPHY, KAY

: APPLICANT: SMITH, TRUDI

: TITLE OF INVENTION: ASP2

: NUMBER OF SEQUENCES: 6

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: RATNER & PRESTIA

: STREET: P.O. BOX 980

: CITY: VALLEY FORGE

: STATE: PA

: COUNTRY: USA

: ZIP: 19482

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FASTSEQ for Windows Version 2.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/009,191

: FILING DATE: 20-JAN-1998

: CLASSIFICATION:

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: UK 9701684.4

: FILING DATE: 28-JAN-1997

: ATTORNEY/AGENT INFORMATION:

: NAME: PRESTIA, PAUL F

: REGISTRATION NUMBER: 23,031

: REFERENCE/DOCKET NUMBER: GH-70368

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 610-407-0700

: TELEFAX: 610-407-0701

: TELEX: 846169

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 2541 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: single

: TOPOLOGY: linear

: MOLECULE TYPE: cDNA

: US-09-009-191-1

Query Match 86.2%; Score 2024; DB 4; Length 2541;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2038; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY	306	ATGCCCAAGCCCTCCCTGGCTCTGCTGTGATGGCGGGGAGTGCTGCTGCCAC	365
DB	1	ATGCCCAAGCCCTCCCTGGCTCTGCTGTGATGGCGGGGAGTGCTGCTGCCAC	60
QY	366	GGCACCAGCAGCGGATCGGCTGCCCTGCGCAGCGGCTGGGGGGCGCCCTGGGG	425
DB	61	GGCACCAGCAGCGGATCGGCTGCCCTGCGCAGCGGCTGGGGGGCGCCCTGGGG	120
QY	426	CTGGCGCTGCCCGGGAGACCGACGAAGACCGCGAGAGCGCGCGCGGAGGCGT	485
DB	121	CTGGCGCTGCCCGGGAGACCGACGAAGACCGCGAGAGCGCGCGCGGAGGCGT	180
QY	486	GTGAGATGGTGACAACTGAGGGCAAGTCGGGCGCAGGCTACTAGTGGAGATGACC	545
DB	181	GTGAGATGGTGACAACTGAGGGCAAGTCGGGCGCAGGCTACTAGTGGAGATGACC	240
QY	546	GTGGCAGCCCCCGCAGACGCTCAACATCCTGGTGATACAGGACGAGTAACCTTGA	605
DB	241	GTGGCAGCCCCCGCAGACGCTCAACATCCTGGTGATACAGGACGAGTAACCTTGA	300

QY	606	GTGGTGCTGCGCCCGCCACCCCTTCTGTCATCTACAGAGCAGCTGTCCAGCACA	665
DB	301	GTGGTGCTGCGCCCGCCACCCCTTCTGTCATCTACAGAGCAGCTGTCCAGCACA	360
QY	666	TACCGGACCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAG	725
DB	361	TACCGGACCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAG	420
QY	726	CTGGCACCAGCTGGTAAGCATCCCCATGGCCCCCAAGCTGCTGCGTCCCAACATT	785
DB	421	CTGGCACCAGCTGGTAAGCATCCCCATGGCCCCCAAGCTGCTGCGTCCCAACATT	480
QY	786	GCTGCCATCACTGAATCAGACAAAGTTCTTTCATCAACGGCTCCAACTGGGAAGGCA	845
DB	481	GCTGCCATCACTGAATCAGACAAAGTTCTTTCATCAACGGCTCCAACTGGGAAGGCA	540
QY	846	GGGTGGGCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTCTTTGACT	905
DB	541	GGGTGGGCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTCTTTGACT	600
QY	906	CTGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTGCAAGCTTTGTGGTCTGGCT	965
DB	601	CTGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTGCAAGCTTTGTGGTCTGGCT	660
QY	966	CCCTCAACCACTCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGGAGTATC	1025
DB	661	CCCTCAACCACTCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGGAGTATC	720
QY	1026	GACCACTGCTGTACACAGGCTGCTGTGTTATACACCCATCCCGGGGAGTGGTAT	1085
DB	721	GACCACTGCTGTACACAGGCTGCTGTGTTATACACCCATCCCGGGGAGTGGTAT	780
QY	1086	GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGACCTGCAAGGAG	1145
DB	781	GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGACCTGCAAGGAG	840
QY	1146	TACAACTATGACAAGAGCATTTGGAGAGTGGACACCAACCTTCTGTTGCCCCAAGAA	1205
DB	841	TACAACTATGACAAGAGCATTTGGAGAGTGGACACCAACCTTCTGTTGCCCCAAGAA	900
QY	1206	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCCTCTCCACGGAGAACTTCCCT	1265
DB	901	GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCCTCTCCACGGAGAACTTCCCT	960
QY	1266	GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGAAGCAGCACCACCTTTGGAACATT	1325
DB	961	GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGAAGCAGCACCACCTTTGGAACATT	1020
QY	1326	TTCCAGTCACTCACTCTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCACTC	1385
DB	1021	TTCCAGTCACTCACTCTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCACTC	1080
QY	1386	ATCCTTCCGACCAATACCTCGGCGCAGTGGGAAGATGGCGCAGTCCCAAGACACTGT	1445
DB	1081	ATCCTTCCGACCAATACCTCGGCGCAGTGGGAAGATGGCGCAGTCCCAAGACACTGT	1140
QY	1446	TACAAGTTTGGCATCTCAGATCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG	1505
DB	1141	TACAAGTTTGGCATCTCAGATCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG	1200
QY	1506	GGCTTCTACGTTGCTTTGATCGGGCCGCAAGAAATGGCTTTGCTGTCAGCGCTTGC	1565
DB	1201	GGCTTCTACGTTGCTTTGATCGGGCCGCAAGAAATGGCTTTGCTGTCAGCGCTTGC	1260
QY	1566	CATGTGCAGGATGAGTTCAGGAGCGGCGGTGGAAGGCCCTTTTGTACCTTGGACATG	1625
DB	1261	CATGTGCAGGATGAGTTCAGGAGCGGCGGTGGAAGGCCCTTTTGTACCTTGGACATG	1320
QY	1626	GAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCCCTCATCACCATTAG	1685
DB	1321	GAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCCCTCATCACCATTAG	1380
QY	1686	GTATGGCTGCCATCTGCGGCCCTCTTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG	1745


```
QY 1326 TTCCAGTCACTCTACTCTACCTAATGGGTGAGGTTTACCAACCAAGTCTTCCGCAATCACC 1385
|
|
|
Db 946 TTCCAGTCACTCTACTCTACCTAATGGGTGAGGTTTACCAACCAAGTCTTCCGCAATCACC 1005
|
|
|
QY 1386 ATCTCTCCGACCAATACCTCGGCCAGTGGAGATGTGGCCAGTCCCAAGACACTGT 1445
|
|
|
Db 1006 ATCTCTCCGACCAATACCTCGGCCAGTGGAGATGTGGCCAGTCCCAAGACACTGT 1065
|
|
|
QY 1446 TACAAGTTTGCCATCTCAGATCATCACGGGCACCTCTTATGGAGCTGTTATCATGAG 1505
|
|
|
Db 1066 TACAAGTTTGCCATCTCAGATCATCACGGGCACCTCTTATGGAGCTGTTATCATGAG 1125
|
|
|
QY 1506 GGCTTCTACGTTGTCTTGTATCGGCCCGGAAACAGAAATGGCTTGTGTGAGCCCTTGC 1565
|
|
|
Db 1126 GGCTTCTACGTTGTCTTGTATCGGCCCGGAAACAGAAATGGCTTGTGTGAGCCCTTGC 1185
|
|
|
QY 1566 CATGTGCAGATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACACCTTGGACATG 1625
|
|
|
Db 1186 CATGTGCAGATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACACCTTGGACATG 1245
|
|
|
QY 1626 GAAGACTGTGCTACAAATTCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
|
|
|
Db 1246 GAAGACTGTGCTACAAATTCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1305
|
|
|
QY 1686 GTCATGTGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCCCTCATGTTGTGAGTGG 1745
|
|
|
Db 1306 GTCATGTGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCCCTCATGTTGTGAGTGG 1365
|
|
|
QY 1746 CGCTGCTCCGCTGCTCGGCCAGACAGATGATGCTTGTGTGATGACATCTCCCTGCTG 1805
|
|
|
Db 1366 CGCTGCTCCGCTGCTCGGCCAGACAGATGATGCTTGTGTGATGACATCTCCCTGCTG 1425
|
|
|
QY 1806 AAGTGAGGAGCCCATGGGCAAGATAGAGATTCCTCGGACACACCTCCGTTGTTCA 1865
|
|
|
Db 1426 AAGTGAGGAGCCCATGGGCAAGATAGAGATTCCTCGGACACACCTCCGTTGTTCA 1485
|
|
|
QY 1866 CTTTGGTGCACAGTAGGAGACAGATGGCACCTGTGGCCAGACCTTCAGGACCCCTCC 1925
|
|
|
Db 1486 CTTTGGTGCACAGTAGGAGACAGATGGCACCTGTGGCCAGACCTTCAGGACCCCTCC 1545
|
|
|
QY 1926 CCACCCACCAATGCTCTGCTTGTATGAGAGAGAAAGGCTGGCAAGTGGGTTCCAG 1985
|
|
|
Db 1546 CCACCCACCAATGCTCTGCTTGTATGAGAGAGAAAGGCTGGCAAGTGGGTTCCAG 1605
|
|
|
QY 1986 GGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAGCACTCTGCTGGCGGGAAATACT 2045
|
|
|
Db 1606 GGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAGCACTCTGCTGGCGGGAAATACT 1665
|
|
|
QY 2046 CTTGGTCACTCAATTTAAGTCGGGAAATTCGTGCTGCTTGAACCTTCAGCCCTGAACT 2105
|
|
|
Db 1666 CTTGGTCACTCAATTTAAGTCGGGAAATTCGTGCTGCTTGAACCTTCAGCCCTGAACT 1725
|
|
|
QY 2106 TTGTCCACCAATTCCTTTAAATTCCTCAACCCCAAGATTCCTTCTTTAGTTTCAGAA 2165
|
|
|
Db 1726 TTGTCCACCAATTCCTTTAAATTCCTCAACCCCAAGATTCCTTCTTTAGTTTCAGAA 1785
|
|
|
QY 2166 GTACTGGCATCACAGCGAGGTACCTTTGGCGTGTGTCCTGTGTGATACCTCGCAGAGAG 2225
|
|
|
Db 1786 GTACTGGCATCACAGCGAGGTACCTTTGGCGTGTGTCCTGTGTGATACCTCGCAGAGAG 1845
|
|
|
QY 2226 AGACCAAGCTTTTCCCTGCTGGCCAAAGTCAAGTAGGAGAGGATGCACAGTTTGTCTATT 2285
|
|
|
Db 1846 AGACCAAGCTTTTCCCTGCTGGCCAAAGTCAAGTAGGAGAGGATGCACAGTTTGTCTATT 1905
|
|
|
QY 2286 TGCTTTAGAGACAGGACTGTATAACACAGCCCTAAACATTTGGTGCAGAGATTTGCCCTTGA 2345
|
|
|
Db 1906 TGCTTTAGAGACAGGACTGTATAACACAGCCCTAAACATTTGGTGCAGAGATTTGCCCTTGA 1965
|
|
|
QY 2346 A 2346
|
|
|
Db 1966 A 1966
```

RESULT 5

US-09-548-367D-5
Sequence 5, Application US/09548367D

Patent No. 6440698

GENERAL INFORMATION:

APPLICANT: GURNEY ET AL.

TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND

TITLE OF INVENTION: THEREOF

FILE REFERENCE: 29915/6280H

CURRENT APPLICATION NUMBER: US/09/548,367D

CURRENT FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: US 60/155,493

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: US 09/404,133

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: PCT/US99/20881

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: US 60/101,594

PRIOR FILING DATE: 1998-09-24

NUMBER OF SEQ ID NOS: 73

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 1977

TYPE: DNA

ORGANISM: Homo sapiens

US-09-548-367D-5

Query Match 80.1%; Score 1881; DB 4; Length 1977;

Best Local Similarity 96.3%; Pred. No. 0;

Matches 1966; Conservative 0; Mismatches 0; Indels 75; Gaps 1;

```
QY 306 ATGGCCCAAGCCCTGCGCTGCTGCTGTGGATGGCGGGAGTGTGCTGCTGCCAC 365
|
|
|
Db 1 ATGGCCCAAGCCCTGCGCTGCTGCTGTGGATGGCGGGAGTGTGCTGCTGCCAC 60
|
|
|
QY 366 GGCACCCAGCACGGCATCCGGCTGCCCTGGCGAGCGGCCCTGGGGGGCCCCCTGGGG 425
|
|
|
Db 61 GGCACCCAGCACGGCATCCGGCTGCCCTGGCGAGCGGCCCTGGGGGGCCCCCTGGGG 120
|
|
|
QY 426 CTGGGCTGCCCGGGGAGACCGAGAGCCCGGAGGAGCCCGGCGGAGGAGGAGGAGCTTT 485
|
|
|
Db 121 CTGGGCTGCCCGGGGAGACCGAGAGCCCGGAGGAGCCCGGCGGAGGAGGAGGAGCTTT 180
|
|
|
QY 486 GTGAGATGGTGACAACTGAGGGGCAAGTGGGGGCAAGTGGGGGCAAGTGGGGGCAAGTGG 545
|
|
|
Db 181 GTGAGATGGTGACAACTGAGGGGCAAGTGGGGGCAAGTGGGGGCAAGTGGGGGCAAGTGG 240
|
|
|
QY 546 GTGGCAGCCCCCGGAGAGCGCTCAACATCTGTGGATACAGGAGGAGTAACTTTGCA 605
|
|
|
Db 241 GTGGCAGCCCCCGGAGAGCGCTCAACATCTGTGGATACAGGAGGAGTAACTTTGCA 300
|
|
|
QY 606 GTGGTGTGCTGCCCGGAGAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 665
|
|
|
Db 301 GTGGTGTGCTGCCCGGAGAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
|
|
|
QY 666 TACCGGAGCTCCCGGAGAGGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
|
|
|
Db 361 TACCGGAGCTCCCGGAGAGGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
|
|
|
QY 726 CTGGGACCGGAGCTGTGTGAGCATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 785
|
|
|
Db 421 CTGGGACCGGAGCTGTGTGAGCATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
|
|
|
QY 786 GCTGCCATCACTGAATCAGACAAAGTCTTCTCAACCGGCTCCAACTGGGAGGAGGAGGAGGAG 845
|
|
|
Db 481 GCTGCCATCACTGAATCAGACAAAGTCTTCTCAACCGGCTCCAACTGGGAGGAGGAGGAGGAG 540
|
|
|
QY 846 GGGCTGGGCTATGCTGAGATTTGCCAGGCTGACGACTCCCTGGAGCCCTTCTTCTTGTGACTCT 905
|
|
|
Db 541 GGGCTGGGCTATGCTGAGATTGCCAG-----566
|
|
|
QY 906 CTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCTCCCTGACGCTTTGTGCTGCTGCTGCTGCT 965
```

Db 567 -----||||| 585
QY 966 CCCCTCAACAGTCTGAAGTGTGGCTCTGTCTGGAGGAGCATGATCATTTGGAGGTATC 1025
Db 586 CCCCTCAACAGTCTGAAGTGTGGCTCTGTCTGGAGGAGCATGATCATTTGGAGGTATC 645
QY 1026 GACCACTCGCTGTACACAGGAGTCTGTGTATACACCCATCGCGGGGAGTGGTATAT 1085
Db 646 GACCACTCGCTGTACACAGGAGTCTGTGTATACACCCATCGCGGGGAGTGGTATAT 705
QY 1086 GAGGTATCATTTGGGGTGGAGATCAATGGACAGATCTGAATAATGGACTGCAAGGAG 1145
Db 706 GAGGTATCATTTGGGGTGGAGATCAATGGACAGATCTGAATAATGGACTGCAAGGAG 765
QY 1146 TACAACATATGACAAGAGCATTTGGGAGTGGGACGACCAACCACTTCCTTTGGCCCAAGAAA 1205
Db 766 TACAACATATGACAAGAGCATTTGGGAGTGGGACGACCAACCACTTCCTTTGGCCCAAGAAA 825
QY 1206 GTGTTTGAAGTGCAGTCAATCAATCAAGGAGGCTCTCCACGGAGAAAGTTCCCTGAT 1265
Db 826 GTGTTTGAAGTGCAGTCAATCAATCAAGGAGGCTCTCCACGGAGAAAGTTCCCTGAT 885
QY 1266 GTTTCCTGAGGAGAGCATTTGGTGTGCTGGCAAGCAGCACCACCCCTTGGAAACAT 1325
Db 886 GTTTCCTGAGGAGAGCATTTGGTGTGCTGGCAAGCAGCACCACCCCTTGGAAACAT 945
QY 1326 TTCCCACTATCTCACTCTACCTAATGGGTGAGGTACCAACCACTCTTCCGCAATCACC 1385
Db 946 TTCCCACTATCTCACTCTACCTAATGGGTGAGGTACCAACCACTCTTCCGCAATCACC 1005
QY 1386 ATCTTTCCGAGCAATACCTCGCGCCAGTGGAGAGTGTGGCCACGTCCTCCCAAGACGACTGT 1445
Db 1006 ATCTTTCCGAGCAATACCTCGCGCCAGTGGAGAGTGTGGCCACGTCCTCCCAAGACGACTGT 1065
QY 1446 TACAAGTTTGGCATCTCACAGTCAATCCACGGGACACTGTATGGAGCTGTATCATGGAG 1505
Db 1066 TACAAGTTTGGCATCTCACAGTCAATCCACGGGACACTGTATGGAGCTGTATCATGGAG 1125
QY 1506 GCGTCTACGTTCTCTTGTATCGGGCCGAGAAACAGAAATGGCTTGTGTGTCAGCGCTTC 1565
Db 1126 GCGTCTACGTTCTCTTGTATCGGGCCGAGAAACAGAAATGGCTTGTGTGTCAGCGCTTC 1185
QY 1566 CATGTGACAGATGAGTTTCAGGAGCGGAGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1625
Db 1186 CATGTGACAGATGAGTTTCAGGAGCGGAGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1245
QY 1626 GAAGACTGTGGCTACACATTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCAT 1685
Db 1246 GAAGACTGTGGCTACACATTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCAT 1305
QY 1686 GTCATGGCTGCCATCTCGCGCCCTTTCATGCTGCCACTCTGCCCTCATGTTGTGTGTCAGTGG 1745
Db 1306 GTCATGGCTGCCATCTCGCGCCCTTTCATGCTGCCACTCTGCCCTCATGTTGTGTGTCAGTGG 1365
QY 1746 CGTGTGCTCGGCTCGGCTCGGCGGAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1805
Db 1366 CGTGTGCTCGGCTCGGCTCGGCGGAGCAGCATGATGACTTTTGTGATGACATCTCCCTGCTG 1425
QY 1806 AAGTGAGAGGCCATCGGCGAGAGATAGAGATTCCTCGGACACACCTCCGTTGGTTCA 1865
Db 1426 AAGTGAGAGGCCATCGGCGAGAGATAGAGATTCCTCGGACACACCTCCGTTGGTTCA 1485
QY 1866 CTTTGTGTACAACTAGGAGACAGATGGCACCTGTGGCCAGAGCAGCTTCAGGACCTCC 1925
Db 1486 CTTTGTGTACAACTAGGAGACAGATGGCACCTGTGGCCAGAGCAGCTTCAGGACCTCC 1545
QY 1926 CCACCCACCAAAATGCTCTGCTGTGATGGAGAGAGAAAGGCTGGCAAGTGGTTCCAG 1985
Db 1546 CCACCCACCAAAATGCTCTGCTGTGATGGAGAGAGAAAGGCTGGCAAGTGGTTCCAG 1605
QY 1986 GGAAGTGTAGGAAACAGAAAAAGAGAAAGAACGACATCTGCTGCGGGAATACT 2045
|||||

Db 1606 GGACTGTACCTGTAGGAAACAGAAAAAGAGAAAGACACTCTGCTGCGGGAATACT 1665
QY 2046 CTTGGTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTCTTCAAACTTTCAGCCCTGAACCT 2105
Db 1666 CTTGGTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTCTTGAACCTTTCAGCCCTGAACCT 1725
QY 2106 TTGTCCACCAATTCCTTTAAATTTCTCCAAACCCAAAGTATTTCTTTCTTTAGTTTTCAGAA 2165
Db 1726 TTGTCCACCAATTCCTTTAAATTTCTCCAAACCCAAAGTATTTCTTTCTTTAGTTTTCAGAA 1785
QY 2166 GTACTGCGATCACACGAGGTTACCTTGGGCTGTGTCCCTGTGTGATACCTTGGGAGAGAG 2225
Db 1786 GTACTGCGATCACACGAGGTTACCTTGGGCTGTGTCCCTGTGTGATACCTTGGGAGAGAG 1845
QY 2226 AGACCAAGCTTGTTCCTCTGCGCCAAAGTTCAGTAGAGAGGATGCACAGTTTGTCTATT 2285
Db 1846 AGACCAAGCTTGTTCCTCTGCGCCAAAGTTCAGTAGAGAGGATGCACAGTTTGTCTATT 1905
QY 2286 TGCTTTAGACAGAGGAGTGTATATAACCAAGCCTTAACATTTGGTGCAAGATTTGCCTCTTGA 2345
Db 1906 TGCTTTAGACAGAGGAGTGTATATAACCAAGCCTTAACATTTGGTGCAAGATTTGCCTCTTGA 1965
QY 2346 A 2346
Db 1966 A 1966

RESULT 6
US-09-009-191-3
: Sequence 3, Application US/09009191
: Patent No. 6319689
: GENERAL INFORMATION:
: APPLICANT: POWELL, DAVID
: APPLICANT: CHAPMAN, CONRAD
: APPLICANT: MURPHY, KAY
: APPLICANT: SMITH, TRUDI
: TITLE OF INVENTION: ASP2
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: RATNER & PRESTIA
: STREET: P.O. BOX 980
: CITY: VALLEY FORGE
: STATE: PA
: COUNTRY: USA
: ZIP: 19482
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/009,191
: FILING DATE: 20-JAN-1998
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: UK 9701684.4
: FILING DATE: 28-JAN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: PRESTIA, PAUL F
: REGISTRATION NUMBER: 23,031
: REFERENCE/DOCKET NUMBER: GH-70368
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-407-0700
: TELEFAX: 610-407-0701
: TELEX: 846169
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2370 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
US-09-009-191-3

Query Match				78.2%;	Score 1835.4;	DB 4;	Length 2370;
Best Local Similarity				99.1%;	Pred. No. 0;		
Matches 1856;				Conservative	0;	Mismatches 16;	Indels 1;
QY	477	GGCAGCTTTGTGGAGATGGTGACAACTGAGGGGCAAGTCGGGGGAGGGCTACTAGCTG	536				
Db	1	GGCAGCTTTGTGGAGATGGTGACAACTGAGGGGCAAGTCGGGGGAGGGCTACTAGCTG	60				
QY	537	GAGATGACCGTGGGACGCCCCCGGAGAGCGTCAACATCCTGTGGATACAGGACGAGT	596				
Db	61	GAGATGACCGTGGGACGCCCCCGGAGAGCGTCAACATCCTGTGGATACAGGACGAGT	120				
QY	597	AACCTTTGAGTGGGTGCTGCCCCCCACCCCTTCTGTCATCGCTACTACAGAGCGAGTG	656				
Db	121	AACCTTTGAGTGGGTGCTGCCCCCCACCCCTTCTGTCATCGCTACTACAGAGAGCGAGTG	180				
QY	657	TCCAGCACATACCGGGACCTCCGGAAGGGTGTATGTGCCCTACACCCAGGCGCAAGTGG	716				
Db	181	TCCAGCACATACCGGGACCTCCGGAAGGGTGTATGAGCCCTACACCCAGGCGCAAGTGG	240				
QY	717	GAAGGGGAGCTGGGACCGACGACCTGTGAAGCATCCCCCATGGCCCCAACGTCACTGTGCGT	776				
Db	241	GAAGGGGAGCTGGGACCGGACGACCTGTGAAGCATCCCCCATGGCCCCAACGTCACTGTGCGT	300				
QY	777	GCCAACTTGTGCCATCACTGAATCAGACAAGTTCTTCATCAACGCTCCAACTGGGAA	836				
Db	301	GCCAACTTGTGCCATCACTGAATCAGACAAGTTCTTCATCAACGCTCCAACTGGGAA	360				
QY	837	GGCATCTGGGGCTGGGCTATGCTGAGATGTCGAGCGCTGACGACCTCCCTGGAGCCTTTC	896				
Db	361	GGCATCTGGGGCTGGGCTATGCTGAGATGTCGAGCGCTGACGACCTCCCTGGAGCCTTTC	420				
QY	897	TTTGACTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAAGCTTTGTGGT	956				
Db	421	TTTGACTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAAGCTTTGTGGT	480				
QY	957	GCTGGCTTCCCTCAACCACTGCTGAAGTGTGGGCTCTGTGCGAGGGAGCATGATCATTT	1016				
Db	481	GCTGGCTTCCCTCAACCACTGCTGAAGTGTGGGCTCTGTGCGAGGGAGCATGATCATTT	540				
QY	1017	GGAGGTATCGACCACTCGCTTACACAGCGAGTCTCTGGTATACCCATCCCGGCGGGAG	1076				
Db	541	GGAGGTATCGACCACTCGCTTACACAGCGAGTCTCTGGTATACCCATCCCGGCGGGAG	600				
QY	1077	TGCTATTATGAGTGTATCTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGAC	1136				
Db	601	TGCTATTATGAGTGTATCTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGAC	660				
QY	1137	TGCAAGGATTAACAATATGACAAGACGATTTGGACAGTGGGACCAACCAACCTTCGTTTG	1196				
Db	661	TGCAAGGATTAACAATATGACAAGACGATTTGGACAGTGGGACCAACCAACCTTCGTTTG	720				
QY	1197	CCCAAGAAAGTGTGAGGCTGCACTAAATCCATCAAGGCGAGCTTCCACGGGAGAAG	1256				
Db	721	CCCAAGAAAGTGTGAGGCTGCACTAAATCCATCAAGGCGAGCTTCCACGGGAGAAG	780				
QY	1257	TTCCCTGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCAACCCCT	1316				
Db	781	TTCCCTGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCAACCCCT	840				
QY	1317	TGGAACATTTTCCAGTCACTCACTTACCTAATGGGTGAGGTTACCAACCAAGTCCCTTC	1376				
Db	841	TGGAACATTTTCCAGTCACTCACTTACCTAATGGGTGAGGTTACCAACCAAGTCCCTTC	900				
QY	1377	CGCATCACCATCTTCCGACGCAATACCTGCGGCCAGTGGGAAGATGCGGCACGTCGCCAA	1436				
Db	901	CGCATCACCATCTTCCGACGCAATACCTGCGGCCAGTGGGAAGATGCGGCACGTCGCCAA	960				
QY	1437	GACGACTGTTACAAGTTTGGCATCTCAGATCATCCAGGGGCACTGTTATGGGAGCTGTT	1496				
Db	961	GACGACTGTTACAAGTTTGGCATCTCAGATCATCCAGGGGCACTGTTATGGGAGCTGTT	1020				

RESULT 7

US-09-548-372D-7
; Sequence 7, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548.372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493

Db 1501 AAGTAAGGAGGCTCGTGGGACAGATGATGGAGACGCCCTGGACACATCTGGGTGGTTCC 1560
Qy 1866 CTTTGGTTCACA--AGTAGGACACAGATGACCTGGCCAGAGCAGCCTCAGGACCT 1923
Db 1561 CTTTGGTTCACATGATGTTGGAGCTATGAGTGTACTGTGGCCAGAGCAGCCTCAGGACCT 1620
Qy 1924 CCCACCCACCAAAATGCTTCGCTTGATGGAGAGGAAAGGCTGGCAAGTGGGTTC 1983
Db 1621 CACCAACTGCAATGCTTCTGGCGTGACAGA-ACAGAGAATCAGGCAAGCTGGATTAC 1679
Qy 1984 AGGACTGTACTGTAGGAACAGAAAAGAGAAGAACTCTGCTGGCGGAATA 2043
Db 1680 AGGCTTGCACCTGTAGGACACAGAGAGGAAGGAAGCAGCGTCTGGTGGCAGGAATA 1739
Qy 2044 CTCCTTGGTCACTCAAAATTAAGTCGGGAATTTCTGCTGCTTGAACCTTCAGCCCTGAAC 2103
Db 1740 TCCTTAGGCACCAAACTTCAGT-TGGAAATTTGCTGCTTGAAGCTTCAGCCCTGACC 1798
Qy 2104 CTTTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGTATTCTTTCTTAGTTTCAG 2163
Db 1799 CTCGTGCCAGCA-TCCTTTAGAGTCTCCAACCTTAAAGTATTCTTTATGTC---CTTCCAG 1854
Qy 2164 AAGTACTGGCATCACACGAGGTTACCTTGGCGGTGTGCTCCCTGGTGGTACCCCTGGCAGAGA 2223
Db 1855 AAGTACTGGCGTCACTACAGGCTACC-CGGCATGTGTCCCTGGTGGTACCCCTGGCAGAGA 1913
Qy 2224 AGACACCAAGCTTGTTCCTCTGCTGGCCAAAGTCACTAGTAGGAGAGATGCACAGTTTGCTA 2283
Db 1914 AAGGCCCAATC-TCATTTCCCTGCTGGCCAAAGTCACTAGTAGGAGAGGTTG--AAGTTTGCCA 1970
Qy 2284 TTTGCTTTAGACAGAGGAGCTGTATAAACAAGCCTTAACATTTGGTGCAGAAAGATTGCTCTT 2343
Db 1971 GTTCTTTAGTAGAGGACTGCAGACTCAAGCCT-ACACTGGTACAAAGAGTGGTCTT 2029
Qy 2344 GA 2345
Db 2030 GA 2031

RESULT 9
US-09-548-372D-29
; Sequence 29, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-372D-29

Query Match 57.7%; Score 1355.8; DB 4; Length 1362;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1357; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 306 ATGGCCCCAAGCCCTGCCCTGCTGCTGATGGCGCGGGAGTGTGCTGCCAC 365
|||||

Db 1 ATGCCCCAAGCCCTGCCCTGCTGCTGATGGCGCGGGAGTGTGCTGCCAC 60
Qy 366 GGCACCCAGCAGCATCCGCTGCCCTGCGCAGCGCTGGGGGGCCCCCTGGG 425
Db 61 GGCACCCAGCAGCATCCGCTGCCCTGCGCAGCGCTGGGGGGCCCCCTGGG 120
Qy 426 CTGGGCTGCCCGGGAGACGACGAGAGCCGAGGAGCCCGCGGAGGGCAGCTTT 485
Db 121 CTGGGCTGCCCGGGAGACGACGAGAGCCGAGGAGCCCGCGGAGGGCAGCTTT 180
Qy 486 GTGAGATGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTAGTGGAGATGACC 545
Db 181 GTGAGATGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTAGTGGAGATGACC 240
Qy 546 GTGGCAGCCCCCGCAGACGCTCAACATCTGCTGGTGGATACAGGCAGCAGTAACCTTTCCA 605
Db 241 GTGGCAGCCCCCGCAGACGCTCAACATCTGCTGGTGGATACAGGCAGCAGTAACCTTTCCA 300
Qy 606 GTGGTGTGCTGCCCGCCACCCCTTCTGCTGCTGCTACTACAGAGGCAGCTGTCCAGCACA 665
Db 301 GTGGTGTGCTGCCCGCCACCCCTTCTGCTGCTGCTACTACAGAGGCAGCTGTCCAGCACA 360
Qy 666 TACGGGACCTCCGGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG 725
Db 361 TACGGGACCTCCGGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG 420
Qy 726 CTGGCAGCCGACCTGGTAAGCATCCCCATGGCCCAACGCTCACTGTGGTGCCCAACATT 785
Db 421 CTGGCAGCCGACCTGGTAAGCATCCCCATGGCCCAACGCTCACTGTGGTGCCCAACATT 480
Qy 786 GCTGCCATCACTGAATCAGACAAGTCTTCTCAACGGCTCAACCTGGGAAGGCATCTCTG 845
Db 481 GCTGCCATCACTGAATCAGACAAGTCTTCTCAACGGCTCAACCTGGGAAGGCATCTCTG 540
Qy 846 GGGCTGGCTATGTGTAGATTCGCCAGGCTGACGACTCCCTGGAGGCTTTCTTTGACTCT 905
Db 541 GGGCTGGCTATGTGTAGATTCGCCAGGCTGACGACTCCCTGGAGGCTTTCTTTGACTCT 600
Qy 906 CTGGTAAAGCAGACCCACGTTCCCAACCTCTCTCCCTGCAGCTTTGTGGTCTGCTGCTC 965
Db 601 CTGGTAAAGCAGACCCACGTTCCCAACCTCTCTCTCCCTGCAGCTTTGTGGTCTGCTGCTC 660
Qy 966 CCCCTCAACACAGTCTGAAGTGTGCGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCCTCAACACAGTCTGAAGTGTGCGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 720
Qy 1026 GACCACTGCTGTACAGGAGGCTGCTGTGTATACACCCATCCCGGGAGTGGTATTAT 1085
Db 721 GACCACTGCTGTACAGGAGGCTGCTGTGTATACACCCATCCCGGGAGTGGTATTAT 780
Qy 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 1145
Db 781 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCAAGGAG 840
Qy 1146 TACAACTATGACAGAGCATTTGTGGACAGTGGCAGCACCACCACTTCTGTTGGCCCAAGAA 1205
Db 841 TACAACTATGACAGAGCATTTGTGGACAGTGGCAGCACCACCACTTCTGTTGGCCCAAGAA 900
Qy 1206 GTGTTTGAAGTGCAGTCAAAATCCATCAAGGAGGCTCTCTCCAGGAGAAAGTTCCTGTAT 1265
Db 901 GTGTTTGAAGTGCAGTCAAAATCCATCAAGGAGGCTCTCTCCAGGAGAAAGTTCCTGTAT 960
Qy 1266 GGTTCCTGCTAGGAGAGCAGCTGCTGCTGAGCAGGAGCAGCACCCTTTGGACATT 1325
Db 961 GGTTCCTGCTAGGAGAGCAGCTGCTGCTGAGCAGGAGCAGCACCCTTTGGACATT 1020
Qy 1326 TTCCCACTCATCTCACTTACCTTAATGGGTGAGGTTACCAACAGTCTCTTCGCGCATCAC 1385
Db 1021 TTCCCACTCATCTCACTTACCTTAATGGGTGAGGTTACCAACAGTCTCTTCGCGCATCAC 1080
Qy 1386 ATCTCTTCCGACAGCAATACCTGCGGGCAGTGGGAAGATGTGGCCACGCTCCCAAGACGACTGT 1445
Db 1081 ATCTCTTCCGACAGCAATACCTGCGGGCAGTGGGAAGATGTGGCCACGCTCCCAAGACGACTGT 1140

Qy	1446	TACAAGTTTGGCATCTCAGATCTACAGGCGACATGTTATGGGAGCTGTTATCATGGAG	1505
Db	1141	TACAAGTTTGGCATCTCAGATCTACAGGCGACATGTTATGGGAGCTGTTATCATGGAG	1200
Qy	1506	GGCTTCTACGTTGCTTTGATCGGGCCGAAACGAATTTGGCTTTCGTCACAGCGTTGC	1565
Db	1201	GGCTTCTACGTTGCTTTGATCGGGCCGAAACGAATTTGGCTTTCGTCACAGCGTTGC	1260
Qy	1566	CATGTGCACGATGAGTTCAGACGCGAGGGTGAAGGCCCTTTTGTGTCACTTTGGACATG	1625
Db	1261	CATGTGCACGATGAGTTCAGACGCGAGGGTGAAGGCCCTTTTGTGTCACTTTGGACATG	1320
Qy	1626	GAAGACTGTGGTACACATTTCCACAGACAGATGAGTCA	1664
Db	1321	GAAGACTGTGGTACACATTTCCACAGACAGATGAGTCA	1359

DEC 11 1964

```

RESULT 10
US-09-548-367D-29
; Sequence 29, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548, 367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-367D-29

```

Query Match	57.7%	Score 1355.8;	DB 4;	Length 1362;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1357	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

Qy	306	ATGCGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGGGAGTGTGCTCTGCCAC	365	Qy	1386	ATCCTTTCCGACGAATACCTGGCGCCAGTGAAGATGTGGCAGCTGCCAAGACGACTGT	1444
Db	1	ATGCGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGGGAGTGTGCTCTGCCAC	60	Db	1081	ATCCTTCCGACGAATACCTGGCGCCAGTGAAGATGTGGCAGCTGCCAAGACGACTGT	1140
Qy	366	GGCACCCAGCAGGGATCGGGGTGCCCCCTGCGAGCGGCTTGGGGGGCCCCCCTGGG	425	Qy	1446	TACAAGTTTCCCATCTCACAGTTCACACGGGCACTGTTATGGAGCTGTTATCATGGAG	1505
Db	61	GGCACCCAGCAGGGATCGGGGTGCCCCCTGCGAGCGGCTTGGGGGGCCCCCCTGGG	120	Db	1141	TACAAGTTTGGCCATCTCACAGTTCACACGGGCACTGTTATGGAGCTGTTATCATGGAG	1200
Qy	426	CTGGGGCTGCCCCGGAGACCCAGCAAGACCCGAGGAGCCCGCCGCGAGGGCAGCTTT	485	Qy	1506	GCCTTCTACGTTCTGCTTTGATCGGGCCCGAAACGAATTTGGCTTTGCTGTCAGCCCTTGC	1565
Db	121	CTGGGGCTGCCCCGGAGACCCAGCAAGACCCGAGGAGCCCGCCGCGAGGGCAGCTTT	180	Db	1201	GCCTTCTACGTTCTGCTTTGATCGGGCCCGAAACGAATTTGGCTTTGCTGTCAGCCCTTGC	1260
Qy	486	GTGGAGATGTGTGGACAACCTGAGGGCAAGTCTGGGGCAGGGCTACTACGTGGAGATGACC	545	Qy	1566	CATGTGCACGATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTTCACCTTGGACATG	1625
Db	181	GTGGAGATGTGTGGACAACCTGAGGGCAAGTCTGGGGCAGGGCTACTACGTGGAGATGACC	240	Db	1261	CATGTGCACGATGAGTTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTTCACCTTGGACATG	1320
Qy	546	GTGGGCAGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACCTTTGCA	605	Qy	1626	GAAGACTGTGGCTTACAACATTTCCACAGACAGATGAGTCA	1664
Db	241	GTGGGCAGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACCTTTGCA	300	Db	1321	GAAGACTGTGGCTTACAACATTTCCACAGACAGATGAGTCA	1359
Qy	606	GTGGGTGTGCCCCGCCACCCCTTCTGCTACGCTACACAGAGGAGCTGTCCAGCACA	665	RESULT 11			
Db	301	GTGGGTGTGCCCCGCCACCCCTTCTGCTACGCTACTACAGAGGAGCTGTCCAGCACA	360	US-09-548-372D-31			
				: Sequence 31, Application US/09548372D			

DEC 11 1971

RESULT II

US-09-548-372D-31

; Sequence 31, Appli

QY	1326	TTCCCAAGTATCTACACTCTTACCTTAATGGGTGAGGTTTACCAACAGTCTCTTCCGCATCACC	1385
Db	1021	TTCCCAAGTATCTACACTCTTACCTTAATGGGTGAGGTTTACCAACAGTCTCTTCCGCATCACC	1080
QY	1386	ATCCCTTCCGAGCAATACCTTCGGCCACTGGAAGATGTGCCACAGTCCCAAGACGACTGT	1445
Db	1081	ATCCCTTCCGAGCAATACCTTCGGCCACTGGAAGATGTGCCACAGTCTCTTCCGCATCACC	1140
QY	1446	TACAAGTTTGGCCATCTCACAGTCTATCCACGGGCACCTGTTTATGGAGCTGTTATCATCGAG	1505
Db	1141	TACAAGTTTGGCCATCTCACAGTCTATCCACGGGCACCTGTTTATGGAGCTGTTATCATCGAG	1200
QY	1506	GGCTTCTACGTTGTCTTTTGATCGGGCCCGAAACAAATTTGGCTTTGCTGTACGCGTTGC	1565
Db	1201	GGCTTCTACGTTGTCTTTTGATCGGGCCCGAAACAAATTTGGCTTTGCTGTACGCGTTGC	1260
QY	1566	CATGTGCACGATGAGTTTCAGGACGGCAGCGGTGGGAAGGCCCTTTTGTCACTTTGGACATG	1625
Db	1261	CATGTGCACGATGAGTTTCAGGACGGCAGCGGTGGGAAGGCCCTTTTGTCACTTTGGACATG	1320
QY	1626	GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCA	1664
Db	1321	GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCA	1359
RESULT 13			
US-09-713-158-1			
; Sequence 1, Application US/09713158			
; Patent No. 6361975			
; GENERAL INFORMATION:			
; APPLICANT: ZHU, YUAN			
; APPLICANT: LI, XIAOTONG			
; APPLICANT: POWELL, DAVID J.			
; APPLICANT: CHRISTIE, GARY			
; TITLE OF INVENTION: MOUSE ASPARTIC SECRETASE-2 (MASP-2)			
; FILE REFERENCE: GP-70660			
; CURRENT APPLICATION NUMBER: US/09713158			
; CURRENT FILING DATE: 2000-11-15			
; PRIOR APPLICATION NUMBER: 60/165,800			
; PRIOR FILING DATE: 1999-11-16			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: FASTSEQ for Windows Version 3.0			
; SEQ ID NO 1			
; LENGTH: 1506			
; TYPE: DNA			
; ORGANISM: MUS MUSCULUS			
US-09-713-158-1			
Query Match 55.3%; Score 1298; DB 4: Length 1506;			
Best Local Similarity 91.4%; Pred. No. 0;			
Matches 1376; Conservative 0; Mismatches 130; Indels 0; Gaps 0;			
QY	306	ATGGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGGCGGGAGTGTGCTGCTGCCAC	365
Db	1	ATGGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGGCGGGAGTGTGCTGCTGCCAC	60
QY	366	GGCACCCAGCACGCGATCCGGCTGCCCTTGGCAGCGGCTTGGGGGGCGCCCTCGGG	425
Db	61	GAACCCATCTCGGCATCCGGCTGCCCTTGCACGCGGCTTGGCAGGGCCACCCCTGGGC	120
QY	426	CTCGCGCTGCCCGGGAGACCGACGAGAGCCGAGAGCCGCGCGGCGAGGGCAGCTTT	485
Db	121	CTGAGGCTGCCCGGGAGACCGACGAGGAATCGGAGAGCCTTGGCCGCGAGAGGACGCTTT	180
QY	486	GTGGAGATGTGGACACACTTAGGGGCAAGTCTGGGGCAGGGCTTACTACGTGGAGATGACC	545
Db	181	GTGGAGATGTGGACACACTTAGGGGCAAGTCTGGGGCAGGGCTTACTACGTGGAGATGACC	240
QY	546	GTGGGAGCCCCCGCAGACGCTCAACATCTCTGTTGGATACAGGCAGCAGTAACCTTCCA	605
Db	241	GTAGGAGCCCCCGCAGACGCTCAACATCTCTGTTGGACACGCGAGTAGTAACCTTCCA	300
QY	606	GTGGGTGCTGCCGCCCAACCCCTTCTGTGATCGCTACTACACAGAGGCAGCTGTGCAGACA	665

Db 301 GTGGGGCTGCCACACACCCCTTTCTGCATCGCTACTACCAGAGGCAAGCTGTCCAGCACA 360
Qy TACCGGACCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAG 725
Db 361 TATCGAGACCTCGGAAGGCTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAA 420
Qy CTGGCACCAGCCTGGTAAGCATCCCCCATGGGCCCAACCTCACTGTGCGTGCCAACATT 785
Db 421 CTGGCACCAGCCTGGTAAGCATCCCCCATGGGCCCAACCTCACTGTGCGTGCCAACATT 480
Qy 786 GTGGCCATCACTGAATCAGACAAGTTCTTCATCAAGGCTCCAACTGGGAAGGCATCCTG 845
Db 481 GCTGCCATCACTGAATCGGACAAGTTCTTCATCAATGTTCCAACTGGGAGGCACTCCTA 540
Qy 846 GGCTGGCTTCTGATGATGGCAGGCTGCAGCACTCCCTGGAGCCTTTCTTTGACTCT 905
Db 541 GGCTGGCTTCTGATGATGGCAGGCTGCAGCACTTTTGGAGCCTTTCTTTGACTCC 600
Qy 906 CTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCTCGCAGCTTTGTGCTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCTCGCAGCTTTGTGCTGGCTTC 660
Qy 966 CCCTCAACCAAGTCTGAAGTCTGGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATC 1025
Db 661 CCCTCAACCAAGCAGCTGGGCTCTGTGCGAGGAGCATGATCATTTGGTGTATC 720
Qy 1026 GACCACTCGCTTACACAGCAGTCTCTGTATACACCCATCCGGCGGAGTGGTATTAT 1085
Db 721 GACCACTCGCTTACAGGCACTCTCTGTATACACCCATCCGGCGGAGTGGTATTAT 780
Qy 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGCATCTGAAATGGACTGCAAGGAG 1145
Db 781 GAAGTATCATTTGACGTGTGGAATCAATGGTCAAGATCTCAAGATGGACTGCAAGGAG 840
Qy 1146 TACAATATGACAGAGCATTGTGGACAGTGGGCACCAACCACTTCCTTTGGCCCAAGAAA 1205
Db 841 TACAATATGACAGAGCATTGTGGACAGTGGGCACCAACCACTTCCTTTGGCCCAAGAAA 900
Qy 1206 GTGTTTGAAGTGCAGTCAATCATCAAGCAGCCTTCCTCCAGGAGAGTTCCCTGAT 1265
Db 901 GTATTTGAAGTGCAGTCAATCATCAAGCAGCCTTCCTCCAGGAGAGTTCCCTGAT 960
Qy 1266 GTTCTTGGCTAGGAGAGCAGCTGGTGTGCGCAAGCAGCAGCAGCCTTTGGAACATT 1325
Db 961 GGCTTTGGCTAGGAGAGCAGCTGGTGTGCGCAAGCAGCAGCAGCCTTTGGAACATT 1020
Qy 1326 TTCCAGTCACTCTACCTAATGGGTGAGTTTACCACCAAGTCTTCCCGCATCACC 1385
Db 1021 TTCCAGTCACTCTACCTAATGGGTGAGTTTACCACCAAGTCTTCCCGCATCACC 1080
Qy 1386 ATCTTTCCGAGCAATACCTGCGGCCAGTGAAGATGTGCCACAGTCCCAAGCAGCTGT 1445
Db 1081 ATCTTTCCGAGCAATACCTGCGGCCAGTGAAGATGTGCCACAGTCCCAAGCAGCTGT 1140
Qy 1446 TACAAGTTTGGCATCTCACAGTATCACAGGCACTGTATGGAGCTGTATATCATCGAG 1505
Db 1141 TACAAGTTTGGCATCTCACAGTATCACAGGCACTGTATGGAGCTGTATATCATCGAG 1200
Qy 1506 GGCTTTACGTTGTCTTTGATCGGGCCGGAAGAAATGGCTTTGCTGTCAGCGTTGC 1565
Db 1201 GGTTCCTATGTCGCTTTCGATCGAGCCCGGAAGGAAATGGCTTTGCTGTCAGCGTTGC 1260
Qy 1566 CATGTGCACATGATTTACAGAGCGGAGCGGTGAAGGCCCTTTGTGACCTTTGGAGATG 1625
Db 1261 CATGTGCACATGATTTACAGAGCGGAGCGGTGAAGGCCCTTTGTGACCTTTGGAGATG 1320
Qy 1626 GAAGACTGTGGCTACAACATTCCACACAGACAGTGAAGTCAACCTTCATGACCATAGCCTAT 1685
Db 1321 GAAGACTGTGGCTACAACATTCCACACAGACAGTGAAGTCAACCTTCATGACCATAGCCTAT 1380
Qy 1686 GTCATGGCTGCCATCTCGGCCCTCTTTCATGCTGCCACTCTGCCTCATGCTGTGTCAGTGG 1745

Db 1381 GTCATGGCGCCATCTGCGCCCTTTCATCTTGCCTATGTCCTATGTCAGTGG 1440
Qy 1746 CGCTGCTCCGCTGCGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGCTCCGCTGCGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1500
Qy 1806 AAGTGA 1811
Db 1501 AAGTGA 1506

RESULT 14

US-09-548-372D-21
; Sequence 21, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-372D-21

Query Match 55.1%; Score 1293.8; DB 4; Length 1341;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 368 CACCCAGCAGCGCATCCGGCTGCCCCCTGCGCAGCGGCTGGGGGGCCCCCCCCCTGGGGCT 427
Db 42 CACCCAGCAGCGCATCCGGCTGCCCCCTGCGCAGCGGCTGGGGGGCCCCCCCCCTGGGGCT 101
Qy 428 CGGCTGCCCGGGAGAGCCAGCAGAGAGCCGAGGAGCCCGGCGGAGGCGACGCTTTGT 487
Db 102 CGGCTGCCCGGGAGAGCCAGCAGAGAGCCCGGAGGAGCCCGGCGGAGGCGACGCTTTGT 161
Qy 488 GGAGATGGTGACAACTGAGGGGCAAGTCCGGGCGAGGCTTACTACGTGGAGATGACCGT 547
Db 162 GGAGATGGTGACAACTGAGGGGCAAGTCCGGGCGAGGCTTACTACGTGGAGATGACCGT 221
Qy 548 GGGCAGCCCCCGGAGAGCGCTCAACATCTCTGGTGATACAGGAGCAGTAACTTTGCAGT 607
Db 222 GGGCAGCCCCCGGAGAGCGCTCAACATCTCTGGTGATACAGGAGCAGTAACTTTGCAGT 281
Qy 608 GGGTGTGCCCCCGGAGAGCGCTTCTGATGCTACTACAGAGGAGCTGTCACAGACATA 667
Db 282 GGGTGTGCCCCCGGAGAGCGCTTCTGATGCTACTACAGAGGAGCTGTCACAGACATA 341
Qy 668 CCGGAGCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAGCT 727
Db 342 CCGGAGCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAGCT 401
Qy 728 GGGCAGCGACTGCTAAAGCATCCCGCATGCCCCCAAGCTACTGCTGGCTGCCAATTCG 787
Db 402 GGGCAGCGACTGCTAAAGCATCCCGCATGCCCCCAAGCTACTGCTGGCTGCCAATTCG 461
Qy 788 TGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCCACTGGGAAGGAGCTCCTGGG 847
Db 462 TGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCCACTGGGAAGGAGCTCCTGGG 521

QY 848 GCTGGCCTATGCTGAGATTGCGAGCCCTGACGACTCCCTGGAGCCCTTTCTTTGACTCTCT 907
Db 522 GCTGGCCTATGCTGAGATTGCGAGCCCTGACGACTCCCTGGAGCCCTTTCTTTGACTCTCT 581
QY 908 GGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGACGCTTTGTGGTGCTGGCTTCCC 967
Db 582 GGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGACGCTTTGTGGTGCTGGCTTCCC 641
QY 968 CCTCAACGAGTCTCAAGTGTGGCTCTGTGGAGGGAGCATGATCATTTGGAGGTATCGA 1027
Db 642 CCTCAACGAGTCTCAAGTGTGGCTCTGTGGAGGGAGCATGATCATTTGGAGGTATCGA 701
QY 1028 CCACTCGCTGTACACAGCAGCTCTGTGTATATACACCCATCCCGGGGAGTGGTATTATGA 1087
Db 702 CCACTCGCTGTACACAGCAGCTCTGTGTATATACACCCATCCCGGGGAGTGGTATTATGA 761
QY 1088 GGTGATCATTTGTGGGTGGAGATCAATGGACAGATCTGAAAATGGACTGCAAGGAGTA 1147
Db 762 GGTGATCATTTGTGGGTGGAGATCAATGGACAGATCTGAAAATGGACTGCAAGGAGTA 821
QY 1148 CAACTATGACAAGAGCATTGTGGAGTGGGACACCAACCTTTCGTTTGGCCCAAGAAAGT 1207
Db 822 CAACTATGACAAGAGCATTGTGGAGTGGGACACCAACCTTTCGTTTGGCCCAAGAAAGT 881
QY 1208 GTTTGAAGCTGCAGTCAAAATCCATCAAGCGAGCCTCTCCACGGAGAAAGTTCCTCTGATGG 1267
Db 882 GTTTGAAGCTGCAGTCAAAATCCATCAAGCGAGCCTCTCCACGGAGAAAGTTCCTCTGATGG 941
QY 1268 TTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGCAGCACCCTTTGGAACATTTT 1327
Db 942 TTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGCAGCACCCTTTGGAACATTTT 1001
QY 1328 CCAGTCTATCTCACTCTACCTTAATGGGTGAGGTACCAACAGTCTCTCCGATCAACCAT 1387
Db 1002 CCAGTCTATCTCACTCTACCTTAATGGGTGAGGTACCAACAGTCTCTCCGATCAACCAT 1061
QY 1388 CTTTCCGACGAATACCTGCGGCGAGTGGAGATGTGGCCAGCTGCCAAGACGACTGTTA 1447
Db 1062 CTTTCCGACGAATACCTGCGGCGAGTGGAGATGTGGCCAGCTGCCAAGACGACTGTTA 1121
QY 1448 CAAGTTTGCCATCTCAGATCATCCAGGCGACTCTTATGGAGCTGTTATCATGAGGG 1507
Db 1122 CAAGTTTGCCATCTCAGATCATCCAGGCGACTCTTATGGAGCTGTTATCATGAGGG 1181
QY 1508 CTTTACGTTGTCTTTGATCGGGCCGAAAACGAATTTGCTTGTGCTGAGCGCTTGCCA 1567
Db 1182 CTTTACGTTGTCTTTGATCGGGCCGAAAACGAATTTGCTTGTGCTGAGCGCTTGCCA 1241
QY 1568 TGTGACGATGAGTTCAGGAGCGGAGCGGTGGAAGGCCCTTTTGTACACCTTGGACATGGA 1627
Db 1242 TGTGACGATGAGTTCAGGAGCGGAGCGGTGGAAGGCCCTTTTGTACACCTTGGACATGGA 1301
QY 1628 AGACTGTGGCTTACAACTTCCACAGACAGATGAGTCA 1664
Db 1302 AGACTGTGGCTTACAACTTCCACAGACAGATGAGTCA 1338

RESULT 15
US-09-548-367D-21
Sequence 21, Application US/09548367D
Patent No. 6440698
GENERAL INFORMATION:
APPLICANT: GURNEY ET AL.
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
FILE REFERENCE: 29915/6280H
CURRENT APPLICATION NUMBER: US/09/548.367D
CURRENT FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: US 60/155,493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 09/404,133
PRIOR FILING DATE: 1999-09-23

QY 368 CACCCAGCAGGCGATCCGGCTGCCCTGCGCAGGCGCTGGGGGCGGCCCTCGGGCT 427
Db 42 CACCCAGCAGGCGATCCGGCTGCCCTGCGCAGGCGCTGGGGGCGGCCCTCGGGCT 101
QY 428 CGCGCTGCCCGGAGACCCACGAAAGAGCCCGAGAGCCCGCGGCGGCGGCGAGGCGCTTTGT 487
Db 102 CGCGCTGCCCGGAGACCCACGAAAGAGCCCGAGAGCCCGCGGCGGCGGCGAGGCGCTTTGT 161
QY 488 GGAGATGTGGACAACCTGAGGGCAAGTGTGGGCGAGGGCTACTACGTGGAGATGACCGT 547
Db 162 GGAGATGTGGACAACCTGAGGGCAAGTGTGGGCGAGGGCTACTACGTGGAGATGACCGT 221
QY 548 GGGCAGCCCGCCGACAGCAGCTCAACATCTCTGCTGATACAGCAGCAGTAACTTTGCAGT 607
Db 222 GGGCAGCCCGCCGACAGCAGCTCAACATCTCTGCTGATACAGCAGCAGTAACTTTGCAGT 281
QY 608 GGGTGTGCTGCCCGCCACCCCTTCTGCTGATCTGCTTACCAGAGGCGAGCTTCTCCAGCACATA 667
Db 282 GGGTGTGCTGCCCGCCACCCCTTCTGCTGATCTGCTTACCAGAGGCGAGCTTCTCCAGCACATA 341
QY 668 CCGGAGCTTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGCAAGTGGGAAGGGAGCT 727
Db 342 CCGGAGCTTCCGGAAGGCTGTATGTGCCCTACACCCAGGCGCAAGTGGGAAGGGAGCT 401
QY 728 GGGCAGCAGCTGCTGAAGCATCCCGCATGCGCCCAACGCTACTGCTGCTGCCAACATTCG 787
Db 402 GGGCAGCAGCTGCTGAAGCATCCCGCATGCGCCCAACGCTACTGCTGCTGCCAACATTCG 461
QY 788 TGCCATCACTGAATCAGACAAAGTCTTTCATCAACGGCTTCCAACTGGGAAGGCGATCCTGGG 847
Db 462 TGCCATCACTGAATCAGACAAAGTCTTTCATCAACGGCTTCCAACTGGGAAGGCGATCCTGGG 521
QY 848 GCTGGCTATGCTGAGATTGCCAGGCTGAGCAGCTCCCTGGAGCCTTTCTTTGACTCTCT 907
Db 522 GCTGGCTATGCTGAGATTGCCAGGCTGAGCAGCTCCCTGGAGCCTTTCTTTGACTCTCT 581
QY 908 GGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGACGCTTTGTGGTGCTGGCTTCCC 967
Db 582 GGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCCTGACGCTTTGTGGTGCTGGCTTCCC 641
QY 968 CTTCAACGAGTCTGAAGTGTGGCTCTGTGGAGGGAGCATGATCATTTGGAGGTATCGA 1027
Db 642 CTTCAACGAGTCTGAAGTGTGGCTCTGTGGAGGGAGCATGATCATTTGGAGGTATCGA 701
QY 1028 CCACCTCGCTGTACACAGGCGATCTCTGATATACACCCATCCCGGGGAGTGGTATTATGA 1087
Db 702 CCACCTCGCTGTACACAGGCGATCTCTGATATACACCCATCCCGGGGAGTGGTATTATGA 761
QY 1088 GGTGATCATTTGTGGGTGGAGATCAATGGACAGATCTGAAAATGGACTGCAAGGAGTA 1147
Db 762 GGTGATCATTTGTGGGTGGAGATCAATGGACAGATCTGAAAATGGACTGCAAGGAGTA 821
QY 1148 CAACTATGACAAGAGCATTGTGGAGTGGGACACCAACCTTTCGTTTGGCCCAAGAAAGT 1207
Db 822 CAACTATGACAAGAGCATTGTGGAGTGGGACACCAACCTTTCGTTTGGCCCAAGAAAGT 881
QY 1208 GTTTGAAGCTGCAGTCAAAATCCATCAAGCGAGCCTCTCCACGGAGAAAGTTCCTCTGATGG 1267

Query Match 55.1%; Score 1293.8; DB 4; Length 1341;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db	882	TTTTGAAGCTGCAGTCAAAATCCATCAAGGAGCGCTCTCCACGGAGAAGTTCCCTGATGG	941
Qy	1268	TTTCTGGCTAGGAGAGCAGCTGGTGTCTGSCAAGCAGGCACCAACCCCTTGGACATTTT	1327
Db	942	TTTCTGGCTAGGAGAGCAGCTGGTGTCTGSCAAGCAGGCACCAACCCCTTGGACATTTT	1001
Qy	1328	CCCAGTCATCTCAGCTACCTACCTAATGGGTGAGGTTACCAACCACTCTTCGCGCATCACCAT	1387
Db	1002	CCCAGTCATCTCAGCTACCTAATGGGTGAGGTTACCAACCACTCTTCGCGCATCACCAT	1061
Qy	1388	CTTTCGCGCAGCAATACCTGCGGCGCAGTGGGAAGATGTGGCCACGTCCCAAGACGACTGTTA	1447
Db	1062	CTTTCGCGCAGCAATACCTGCGGCGCAGTGGGAAGATGTGGCCACGTCCCAAGACGACTGTTA	1121
Qy	1448	CAAGTTTGGCCATCTCACAGTCATCCACGGGCACCTGTTATGGAGCTCTTATCATGGAGGG	1507
Db	1122	CAAGTTTGGCCATCTCACAGTCATCCACGGGCACCTGTTATGGAGCTCTTATCATGGAGGG	1181
Qy	1508	CTTCTACGTTGTCTTTGATCGGGCCCGAAACGAATTTGGCTTTGCTGTACGCGCTTGCCA	1567
Db	1182	CTTCTACGTTGTCTTTGATCGGGCCCGAAACGAATTTGGCTTTGCTGTACGCGCTTGCCA	1241
Qy	1568	TGTGCGCAGATGAGTTCAGGAGCGGCGGTGGGAAGGCCCTTTTGTACCTTGGACATGGA	1627
Db	1242	TGTGCGCAGATGAGTTCAGGAGCGGCGGTGGGAAGGCCCTTTTGTACCTTGGACATGGA	1301
Qy	1628	AGACTGTGGCTACAACATTCACAGACAGATGAGTCA	1664
Db	1302	AGACTGTGGCTACAACATTCACAGACAGATGAGTCA	1338

Search completed: March 2, 2003, 06:16:13
Job time : 76 secs

QY 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAAAATGACTGCAAGGAG 1145
|||||
Db 781 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGGATCTGAAAAATGACTGCAAGGAG 840
QY 1146 TACAACATATGACAAAGAGCATTTGGAGAGTGGGACCAACCACTTCGTTGGCCCAAGAAA 1205
|||||
Db 841 TACAACATATGACAAAGAGCATTTGGAGAGTGGGACCAACCACTTCGTTGGCCCAAGAAA 900
QY 1206 GTGTTTGAAGCTGCAGTCAATCCATCAAGGCAGCCTCCTCCACGGGAGAAGTTCCCTGAT 1265
|||||
Db 901 GTGTTTGAAGCTGCAGTCAATCCATCAAGGCAGCCTCCTCCACGGGAGAAGTTCCCTGAT 960
QY 1266 GGTTCCTGCTAGGAGAGCAGTGGTGTGCTGGCAAGCAGGCACCAACCCCTTGGAAACATT 1325
|||||
Db 961 GGTTCCTGCTAGGAGAGCAGTGGTGTGCTGGCAAGCAGGCACCAACCCCTTGGAAACATT 1020
QY 1326 TTCCAGTCACTCTACTCTACCTAATAGGTGAGGTTACCAACCAAGTCCCTCCGCATCAC 1385
|||||
Db 1021 TTCCAGTCACTCTACTCTACCTAATAGGTGAGGTTACCAACCAAGTCCCTCCGCATCAC 1080
QY 1386 ATCTTCCGACGATACCTGCGGCCAGTGGAAAGATGTGGCCAGCTCCCAAGACGACTGT 1445
|||||
Db 1081 ATCTTCCGACGATACCTGCGGCCAGTGGAAAGATGTGGCCAGCTCCCAAGACGACTGT 1140
QY 1446 TACAAGTTTGGCATCTCAGATCATCCACGGGCACCTGTTATGGAGCTGTTATCATGGAG 1505
|||||
Db 1141 TACAAGTTTGGCATCTCAGATCATCCACGGGCACCTGTTATGGAGCTGTTATCATGGAG 1200
QY 1506 GGCTTCTACGTTGCTTTGATCGGGCCGGAACGAATTTGGCTTGTGTCAGCGCTTGC 1565
|||||
Db 1201 GGCTTCTACGTTGCTTTGATCGGGCCGGAACGAATTTGGCTTGTGTCAGCGCTTGC 1260
QY 1566 CATGTGACGATAGTTTACGAGCGGCGGTGGAGGCCCTTTTGTACCTTTGACATG 1625
|||||
Db 1261 CATGTGACGATAGTTTACGAGCGGCGGTGGAGGCCCTTTTGTACCTTTGACATG 1320
QY 1626 GAAGACTGTGGCTACAACATTTCCACAGACGATGAGTCAACCCCTCATGACCATACGCTAT 1685
|||||
Db 1321 GAAGACTGTGGCTACAACATTTCCACAGACGATGAGTCAACCCCTCATGACCATACGCTAT 1380
QY 1686 GTCATGGCTGCGATCTGCGCCCTCTTCATGCTGCGCACTGCGCCATGCTGTCAGTGG 1745
|||||
Db 1381 GTCATGGCTGCGATCTGCGCCCTCTTCATGCTGCGCACTGCGCCATGCTGTCAGTGG 1440
QY 1746 CGCTGCTCCGCTGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1805
|||||
Db 1441 CGCTGCTCCGCTGCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGAGAGGCCCATGGSCAGAAGATAGAGATTCCTCCCTGGACACACCTCCGTTGTTCA 1865
|||||
Db 1501 AAGTGAGAGGCCCATGGSCAGAAGATAGAGATTCCTCCCTGGACACACCTCCGTTGTTCA 1560
QY 1866 CTTTGGTGCACAGTAGGAGACACAGATGGACCTGTGGCCAGAGCACCTCAGGACCCCTCC 1925
|||||
Db 1561 CTTTGGTGCACAGTAGGAGACACAGATGGACCTGTGGCCAGAGCACCTCAGGACCCCTCC 1620
QY 1926 CCACCCACCAATGCTCTGCTCTTATGGAGAGGAAAGGCTGGCAAGGTTGGGTTCCAG 1985
|||||
Db 1621 CCACCCACCAATGCTCTGCTCTTATGGAGAGGAAAGGCTGGCAAGGTTGGGTTCCAG 1680
QY 1986 GGACTGTACCTGTAGGAAACAGAAAAGAGAAAGAGCACTCTGCTGGGGGGAATACT 2045
|||||
Db 1681 GGACTGTACCTGTAGGAAACAGAAAAGAGAAAGAGCACTCTGCTGGGGGGAATACT 1740
QY 2046 CTTGGTCACTCAAAATTTAAGTCGGGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2105
|||||
Db 1741 CTTGGTCACTCAAAATTTAAGTCGGGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY 2106 TTGTCCACCATTCCTTTAAATTTCTCCAAACCAAGATTTCTTCTTTTCTTTTCTTTTCTTTT 2165
|||||
Db 1801 TTGTCCACCATTCCTTTAAATTTCTCCAAACCAAGATTTCTTCTTTTCTTTTCTTTTCTTTT 1860

QY 2166 GTACTGGCATCATCACAGCAGGTTACCTTTGGCTGTGTCCTCTGTGGTACCTTGGCAGAG 2225
|||||
Db 1861 GTACTGGCATCATCACAGCAGGTTACCTTTGGCTGTGTCCTCTGTGGTACCTTGGCAGAG 1920
QY 2226 AGACCAAGCTTTGTTTCCCTGCTGCGCAAAAGTTCAGTAGGAGAGGATTCACAGTTTGCATT 2285
|||||
Db 1921 AGACCAAGCTTTGTTTCCCTGCTGCGCAAAAGTTCAGTAGGAGAGGATTCACAGTTTGCATT 1980
QY 2286 TGCTTTAGAGAGAGGACCTGTATAAACAAGCCTAACATTGGTGCAGAAAGATTGCTCTTGA 2345
|||||
Db 1981 TGCTTTAGAGAGAGGACCTGTATAAACAAGCCTAACATTGGTGCAGAAAGATTGCTCTTGA 2040
QY 2346 ATT 2348
|||||
Db 2041 ATT 2043
RESULT 5
US-09-794-925-3
: Sequence 3, Application US/09794925
: Patent No. US20020064819A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Bienkowski, Michael J.
: APPLICANT: Heinrikson, Robert L.
: APPLICANT: Parodi, Luis A.
: APPLICANT: Yan, Riqiang
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 28341/6280H1
: CURRENT APPLICATION NUMBER: US/09/794,925
: CURRENT FILING DATE: 2001-02-27
: PRIOR APPLICATION NUMBER: 09/416,901
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: 60/155,493
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 09/404,133
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: PCT/US99/20881
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 60/101,594
: PRIOR FILING DATE: 1998-09-24
: NUMBER OF SEQ ID NOS: 73
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 2070
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-794-925-3
Query Match 86.9%; Score 2039.8; DB 10; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 306 ATGGCCCAAGCCCTGCGCTGCTGCTGTGGATGGCGGGGAGTGTGCTGCTGCCAC 365
|||||
Db 1 ATGGCCCAAGCCCTGCGCTGCTGCTGTGGATGGCGGGGAGTGTGCTGCTGCCAC 60
QY 366 GGCACCCACGACGGCATCCGGCTGCCCCCTGCGCAGCGCCCTGGGGGGCCCCCTGGGG 425
|||||
Db 61 GGCACCCACGACGGCATCCGGCTGCCCCCTGCGCAGCGCCCTGGGGGGCCCCCTGGGG 120
QY 426 CTGGGCTCCCCGGGAGACCGAGAGAGCCGAGAGCCGCGGAGGCGGCGGAGGCGAGCTTT 485
|||||
Db 121 CTGGGCTCCCCGGGAGACCGAGAGAGCCGAGAGCCGCGGAGGCGGCGGAGGCGAGCTTT 180
QY 486 GTGGAGATGGTGACAACTGAGGGGCAAGTCTGGGGCAGGCGCTACTACGTGGAGATGACC 545
|||||
Db 181 GTGGAGATGGTGACAACTGAGGGGCAAGTCTGGGGCAGGCGCTACTACGTGGAGATGACC 240
QY 546 GTGGCAGCCCCCGCAGACGCTCAACATCTGTGGATACAGGAGCAGTAACCTTTGCA 605
|||||
Db 241 GTGGCAGCCCCCGCAGACGCTCAACATCTGTGGATACAGGAGCAGTAACCTTTGCA 300

Db 1501 AAGTGAGAGGCCCCATGGAGAGAAAGATAGAGATTCCCTCGGACACACACCTCCGTGGTTC 1560
QY 1865 ACTTTGGTCAACAAGTAGGAGACACAGATGCCACCTGTGGCCAGAGCAGCCTCAGGACCCCTC 1924
Db 1561 ACTTTGGTCAACAAGTAGGAGACACAGATGGCAGCTGTGGCCAGAGCAGCCTCAGGACCCCTC 1620
QY 1925 CCCACCCACCAAAATGCCCTCTGCTTGTGATGGAGAGGAAAGGCTGGCAAGTGGTTCCA 1984
Db 1621 CCCACCCACCAAAATGCCCTCTGCTTGTGATGGAGAGGAAAGGCTGGCAAGTGGTTCCA 1680
QY 1985 GGGAGCTGTACCTGTAGGAACACAGAAAGAGAAAGAACACACCTCTGCTGGCGGAATAC 2044
Db 1681 GGGAGCTGTACCTGTAGGAACACAGAAAGAGAAAGAACACCTCTGCTGGCGGAATAC 1740
QY 2045 TCTTGGTCACCTCAAAATTAAGTCGGGAAATCTGCTGCTTGAACCTCAGCCCTGAACC 2104
Db 1741 TCTTGGTCACCTCAAAATTAAGTCGGGAAATCTGCTGCTTGAACCTCAGCCCTGAACC 1800
QY 2105 TTTGTCACCAATCTTTAAATTCCTCAACCCCAAGATATCTTCTTTTCTTAGTTTCAGA 2164
Db 1801 TTTGTCACCAATCTTTAAATTCCTCAACCCCAAGATATCTTCTTTTCTTAGTTTCAGA 1860
QY 2165 AGTACTGSCATCACACGAGGTTACCTTGGCGTGTGTCCTGTGCTACCTGGCAGAGAA 2224
Db 1861 AGTACTGSCATCACACGAGGTTACCTTGGCGTGTGTCCTGTGCTACCTGGCAGAGAA 1920
QY 2225 GAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGCTAT 2284
Db 1921 GAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGCTAT 1980
QY 2285 TTGCTTTAGACAGGAGCTGTATAAAGCCATAACATTTGGTGCAAGATTGCTCTTGG 2344
Db 1981 TTGCTTTAGACAGGAGCTGTATAAAGCCATAACATTTGGTGCAAGATTGCTCTTGG 2040
QY 2345 AATT 2348
Db 2041 AATT 2044

RESULT 8
US-09-795-903A-1
; Sequence 1, Application US/09795903A
; Patent No. US20020164760A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
; FILE REFERENCE: of Use Thereof
; CURRENT APPLICATION NUMBER: US/09/795,903A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/604,608
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-903A-1

Query Match 85.3%; Score 2004; DB 9; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 GCGGAGTGTCTGCTGCCACAGCGCACCCAGCACGATCCCGCTCCCTTGCAGCGGC 404
Db 1 GCGGAGTGTCTGCTGCCACAGCGCACCCAGCACGATCCCGCTCCCTTGCAGCGGC 60
QY 405 CTGGGGGGGGCCCCCTGGGGCTGGCGCTGCCCGGAGAGCCGAGAGAGAGAGAGAG 464
Db 61 CTGGGGGGGGCCCCCTGGGGCTGGCGCTGCCCGGAGAGCCGAGAGAGAGAGAGAG 120
QY 465 CCGGCGGAGGGGAGCTTTGTGGAGATGCTGACACACCTCAGGGGGAAGTCGGGGCAG 524
Db 121 CCGGCGGAGGGGAGCTTTGTGGAGATGCTGACACACCTCAGGGGGAAGTCGGGGCAG 180
QY 525 GGCTACTACGTGAGATGACCGTGGGAGCGCCCGGAGAGCGCTCAACATCTCTGGTAT 584
Db 181 GGCTACTACGTGAGATGACCGTGGGAGCGCCCGGAGAGCGCTCAACATCTCTGGTAT 240
QY 585 ACAGCAGCAGTAACCTTTGAGTGGGTGCTGCCCGCCACCCCTTCCCTGCATCGCTACTAC 644
Db 241 ACAGCAGCAGTAACCTTTGAGTGGGTGCTGCCCGCCACCCCTTCCCTGCATCGCTACTAC 300
QY 645 CAGAGCGAGCTCTCCAGCACATACCGGAGCTCCCGGAGGCTGCTATGTCGCCCTACACC 704
Db 301 CAGAGCGAGCTCTCCAGCACATACCGGAGCTCCCGGAGGCTGCTATGTCGCCCTACACC 360
QY 705 CAGGCAAGTGGGAGGGAGCTGGGACGACCTGGTAAAGCATCCCGCATGGCCCCAAC 764
Db 361 CAGGCAAGTGGGAGGGAGCTGGGACGACCTGGTAAAGCATCCCGCATGGCCCCAAC 420
QY 765 GTCACTGTGCTGCCAACATTTGCTGCCATCAGTCAATCAGACAGTTCTTCAATCAACGGC 824
Db 421 GTCACTGTGCTGCCAACATTTGCTGCCATCAGTCAATCAGACAGTTCTTCAATCAACGGC 480
QY 825 TCCAACTGGGAGGAGCTCTGGGCTGAGTATGCTGAGATGGCAGGCTCAGCAGCTCC 884
Db 481 TCCAACTGGGAGGAGCTCTGGGCTGAGTATGCTGAGATGGCAGGCTCAGCAGCTCC 540
QY 885 CTGGAGCCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTG 944
Db 541 CTGGAGCCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACCTTCCCAACCTCTTCTCCCTG 600
QY 945 CAGCTTTGTGCTGGCTTCCCGCTCAACAGCTCTGAAGTGTGGGCTCTCTCGGAGGG 1004
Db 601 CAGCTTTGTGCTGGCTTCCCGCTCAACAGCTCTGAAGTGTGGGCTCTCTCGGAGGG 660
QY 1005 AGCATGATCATTTGGAGGTATCGACCTCGCTGTACAGAGCAGTCTCTGGTATACACC 1064
Db 661 AGCATGATCATTTGGAGGTATCGACCTCGCTGTACAGAGCAGTCTCTGGTATACACC 720
QY 1065 ATCCGGCGGGAGTGGTATTTATGAGTGTATCTTGGGGTGGAGATCAATGGACAGGAT 1124
Db 721 ATCCGGCGGGAGTGGTATTTATGAGTGTATCTTGGGGTGGAGATCAATGGACAGGAT 780
QY 1125 CTGAAAATGGAGTCAAGAGTACAACTATGACAAGAGCATTTGGGAGAGTGGCACCACC 1184
Db 781 CTGAAAATGGAGTCAAGAGTACAACTATGACAAGAGCATTTGGGAGAGTGGCACCACC 840
QY 1185 AACCTTCGTTTCCCAAGAAAGTGTGAAAGTCAAGTCAATCCATCAAGCAGCCTCC 1244
Db 841 AACCTTCGTTTCCCAAGAAAGTGTGAAAGTCAAGTCAATCCATCAAGCAGCCTCC 900
QY 1245 TCCAGCGGAGAGTTCCTCTGATGGTTCCTGCTAGGAGAGCAGTGGTGTCTGGCAAGCA 1304
Db 901 TCCAGCGGAGAGTTCCTCTGATGGTTCCTGCTAGGAGAGCAGTGGTGTCTGGCAAGCA 960
QY 1305 GGCACCAACCTTTGGAACATTTTCCAGTCTATCTACCTCTACCTTAATGGGTGAGGTTACC 1364
Db 961 GGCACCAACCTTTGGAACATTTTCCAGTCTATCTACCTCTACCTTAATGGGTGAGGTTACC 1020
QY 1365 AACCACTCTTCCCGCATCACCTCTTCCAGCAGCAATACCTTGGGCGCAGTGGCAAGATGTG 1424
Db 1021 AACCACTCTTCCCGCATCACCTCTTCCAGCAGCAATACCTTGGGCGCAGTGGCAAGATGTG 1080

Qy 1425 GCCACGTCCTCCAGAGGACTGTTACAAAGTTTCCATCTCACAGTCATCCACGGGCACTGTT 1484
Db 1081 GCCAGTCCCAAGAGGACTGTTACAAAGTTTCCATCTCACAGTCATCCACGGGCACTGTT 1140
Qy 1485 ATGGAGCTGTTATCATGAGGAGCTTACCTGTTGCTTTGATCGGGCCCGAAACGAATT 1544
Db 1141 ATGGAGCTGTTATCATGAGGAGCTTACCTGTTGCTTTGATCGGGCCCGAAACGAATT 1200
Qy 1545 GGCCTTCTGTCAGCGCTTGCATGTCACCATGAGTTCAGAGCGGACGGTGGAGGC 1604
Db 1201 GGCCTTCTGTCAGCGCTTGCATGTCACCATGAGTTCAGAGCGGACGGTGGAGGC 1260
Qy 1605 CCTTTTGTACCTTGGACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCA 1664
Db 1261 CCTTTTGTACCTTGGACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCA 1320
Qy 1665 ACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTGGCGCCCTTTCATGCTGCCACTC 1724
Db 1321 ACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTGGCGCCCTTTCATGCTGCCACTC 1380
Qy 1725 TGCCTCATGTTGTCAGTGGCGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCT 1784
Db 1381 TGCCTCATGTTGTCAGTGGCGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCT 1440
Qy 1785 GCTGATGACATCTCCCTGCTGAACTGAGGAGGCGCCATGGGAGAGATAGAGATTCCCGCT 1844
Db 1441 GCTGATGACATCTCCCTGCTGAACTGAGGAGGCGCCATGGGAGAGATAGAGATTCCCGCT 1500
Qy 1845 GGACACACCTCGTGGTTCACCTTGTGTCACAACTAGGAGACACAGATGGCACCTGTGGC 1904
Db 1501 GGACACACCTCGTGGTTCACCTTGTGTCACAACTAGGAGACACAGATGGCACCTGTGGC 1560
Qy 1905 CAGAGCACTTCAGAGCCCTCCACCCACCAACCAATGCCCTGCCCTGATGGAGAGGAAAA 1964
Db 1561 CAGAGCACTTCAGAGCCCTCCACCCACCAACCAATGCCCTGCCCTGATGGAGAGGAAAA 1620
Qy 1965 GGCTGGCAAGTGGGTTCCAGGGACTGTACCTGTAGGAAACAGAAAAAGAGAAAGAAAG 2024
Db 1621 GGCTGGCAAGTGGGTTCCAGGGACTGTACCTGTAGGAAACAGAAAAAGAGAAAGAAAG 1680
Qy 2025 CACTCTGCTGGGGGAATCTCTTGGTCACTCACTCAAAATTAAGTCGGGAAATTCCTGCT 2084
Db 1681 CACTCTGCTGGGGGAATCTCTTGGTCACTCACTCAAAATTAAGTCGGGAAATTCCTGCT 1740
Qy 2085 TGAACATTCAGCCCTGAACTTTGTCCACCACTTCTTTAAATTCCTCAACCCAAAGTATT 2144
Db 1741 TGAACATTCAGCCCTGAACTTTGTCCACCACTTCTTTAAATTCCTCAACCCAAAGTATT 1800
Qy 2145 CTTCTTTCTTTAGTTTCAGAACTACTGGCATCACACGAGGTTACCTTGGCGTGTGCTCC 2204
Db 1801 CTTCTTTCTTTAGTTTCAGAACTACTGGCATCACACGAGGTTACCTTGGCGTGTGCTCC 1860
Qy 2205 TGTGTTACCTTGGCAGAGAGAGACCAAGCTTGTTCCTGCTGCCCAAGTCAGTAGGA 2264
Db 1861 TGTGTTACCTTGGCAGAGAGAGACCAAGCTTGTTCCTGCTGCCCAAGTCAGTAGGA 1920
Qy 2265 CAGGATGACAGATTGCTTATTTGCTTTAGACACAGGGACTGTATAAACAAAGCCTAAACATT 2324
Db 1921 CAGGATGACAGATTGCTTATTTGCTTTAGACACAGGGACTGTATAAACAAAGCCTAAACATT 1980
Qy 2325 GGTGCAAGATTGGCTCTTTGAATT 2348
Db 1981 GGTGCAAGATTGGCTCTTTGAATT 2004

RESULT 9

US-09-796-264-1
: Sequence 1, Application US/09796264
: Patent No. US20020049303A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Jordan J.N.
: APPLICANT: Lin, Xinli
: APPLICANT: Koelsch, Gerald

: TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
: TITLE OF INVENTION: of use Thereof ;
: FILE REFERENCE: OMRF 179
: CURRENT APPLICATION NUMBER: US/09/796,264
: PRIOR FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: 09/604,608
: PRIOR FILING DATE: 2000-06-27
: PRIOR APPLICATION NUMBER: 60/168,060
: PRIOR FILING DATE: 1999-11-30
: PRIOR APPLICATION NUMBER: 60/177,836
: PRIOR FILING DATE: 2000-01-25
: PRIOR APPLICATION NUMBER: 60/178,368
: PRIOR FILING DATE: 2000-01-27
: PRIOR APPLICATION NUMBER: 60/210,292
: PRIOR FILING DATE: 2000-06-08
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 3252
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-796-264-1

Query Match 85.3%; Score 2004; DB 10; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 345 CGGGAGTGTGCTGCCACGGCACCCAGCAGCGATCCGGCTGCCCTGGCGAGCGGC 404
Db 1 CGGGAGTGTGCTGCCACGGCACCCAGCAGCGATCCGGCTGCCCTGGCGAGCGGC 60
Qy 405 CTGGGGCGCGCCCTCGGGCTGGGATGGTGGAGATGGTGGAGAACCTTGAGGGGCAAGTCGGGGCAG 464
Db 61 CTGGGGCGCGCCCTCGGGCTGGGATGGTGGAGATGGTGGAGAACCTTGAGGGGCAAGTCGGGGCAG 120
Qy 465 CCCGGCGAGGGGAGCTTTGTGGAGATGGTGGAGAACCTTGAGGGGCAAGTCGGGGCAG 524
Db 121 CCCGGCGAGGGGAGCTTTGTGGAGATGGTGGAGAACCTTGAGGGGCAAGTCGGGGCAG 180
Qy 525 GGCTACTACGTGGAGATGACCGTGGGAGCGCCCGGAGAGCGCTCAACATCTGCTGGAT 584
Db 181 GGCTACTACGTGGAGATGACCGTGGGAGCGCCCGGAGAGCGCTCAACATCTGCTGGAT 240
Qy 585 ACAGCAGCAGTAACTTTGTCAGTGGTGTGCCCGCCCGGAGAGCGCTTCTGCTGCTACTAC 644
Db 241 ACAGCAGCAGTAACTTTGTCAGTGGTGTGCCCGCCCGGAGAGCGCTTCTGCTGCTACTAC 300
Qy 645 CAGAGGAGCTGTCCAGCACATACCGGAGCTTCCGGAAGGTTGTATGTGCCCTTACACC 704
Db 301 CAGAGGAGCTGTCCAGCACATACCGGAGCTTCCGGAAGGTTGTATGTGCCCTTACACC 360
Qy 705 CAGGCAAGTGGAGGGAGCTGGGACCGGAGCGCTTAAAGCATCCCGCATGGCCCCAAC 764
Db 361 CAGGCAAGTGGAGGGAGCTGGGACCGGAGCGCTTAAAGCATCCCGCATGGCCCCAAC 420
Qy 765 GTCATGTGCTGCCCAACATTTGCTGCCATCACTGAATCAGACAAGTTCTTCAACAGCG 824
Db 421 GTCATGTGCTGCCCAACATTTGCTGCCATCACTGAATCAGACAAGTTCTTCAACAGCG 480
Qy 825 TCCAACTGGGAAGGATCCTGGGGCTGGCCTATGCTGAGATGGCAGGCTGACGACTCC 884
Db 481 TCCAACTGGGAAGGATCCTGGGGCTGGCCTATGCTGAGATGGCAGGCTGACGACTCC 540
Qy 885 CTGGAGCGCTTCTTGTGACTCTCTGTAAGCAGAGCGGCTTCCCAACCTTCTTCCCTG 944
Db 541 CTGGAGCGCTTCTTGTGACTCTCTGTAAGCAGAGCGGCTTCCCAACCTTCTTCCCTG 600
Qy 945 CAGCTTTTGTGCTGGCTTCCCGCTCAACAGTCTGAAGTGTGGCTTCTCTCGGAGG 1004
Db 601 CAGCTTTTGTGCTGGCTTCCCGCTCAACAGTCTGAAGTGTGGCTTCTCTCGGAGG 560
Qy 1005 AGCATGATCTTGGAGGTATCGACCACTCGCTGTATACAGAGGAGCTCTCTGTTATACACC 1064
Db 1005 AGCATGATCTTGGAGGTATCGACCACTCGCTGTATACAGAGGAGCTCTCTGTTATACACC 1064

Db 661 AGCATGATCATGGAGGTATCGACCACTCGCTGTACACAGGCACTCTCTGGTATACACCC 720
QY 1065 ATCCGGCGGGAGTGTATATGAGTGTGATCATGTGCGGGTGGAGATCAATGGACAGGAT 1124
Db 721 ATCCGGCGGGAGTGTATATGAGTGTGATCATGTGCGGGTGGAGATCAATGGACAGGAT 780
QY 1125 CTGAATATGGACTGCAAGAGTACAACTATGACAAAGAGATTTGGACAGTGGCAACCACC 1184
Db 781 CTGAATATGGACTGCAAGAGTACAACTATGACAAAGAGATTTGGACAGTGGCAACCACC 840
QY 1185 AACCTTCCTTTGCCCAAGAAAGTGTGTAAGCTGCGAGTCAAAATCCATCAAGSGACGCTCC 1244
Db 841 AACCTTCCTTTGCCCAAGAAAGTGTGTAAGCTGCGAGTCAAAATCCATCAAGSGACGCTCC 900
QY 1245 TCCACGGAGAAAGTTCCTCGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 1304
Db 901 TCCACGGAGAAAGTTCCTCGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 960
QY 1305 GGCACCAACCTTGAACAAATTTCCAGTCACTCTCACTTAATGGGTGAGGTTACC 1364
Db 961 GGCACCAACCTTGAACAAATTTCCAGTCACTCTCACTTAATGGGTGAGGTTACC 1020
QY 1365 AACCAGTCTTCCCGCATCACCATCTTCCGCGAGCAATACCTGCGCCAGTGGAGATGTG 1424
Db 1021 AACCAGTCTTCCCGCATCACCATCTTCCGCGAGCAATACCTGCGCCAGTGGAGATGTG 1080
QY 1425 GCCACGTCACAGAGTGTGTACAAAGTTTGCCATCTCACAGTCACTCCACGGGCACTGT 1484
Db 1081 GCCACGTCACAGAGTGTGTGTACAAAGTTTGCCATCTCACAGTCACTCCACGGGCACTGT 1140
QY 1485 ATGGAGCTGTATCATGAGGGCTTCTAGTGTGTCTTGTATCGGGCCCGAAACGAAT 1544
Db 1141 ATGGAGCTGTATCATGAGGGCTTCTAGTGTGTCTTGTATCGGGCCCGAAACGAAT 1200
QY 1545 GGCCTTGTCTCAGCGCTTGCCATGTGCACGATGTGCTAGGAGCGGCGGTGGAGGC 1604
Db 1201 GGCCTTGTCTCAGCGCTTGCCATGTGCACGATGTGCTAGGAGCGGCGGTGGAGGC 1260
QY 1605 CTTTGTGACCTTGGACATGGAAGACTGTGGCTACAACATTCACACAGACAGATGAGTCA 1664
Db 1261 CTTTGTGACCTTGGACATGGAAGACTGTGGCTACAACATTCACACAGACAGATGAGTCA 1320
QY 1665 ACCCTCATGACATAGCCTATGCTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1724
Db 1321 ACCCTCATGACATAGCCTATGCTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1725 TGCTCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1784
Db 1381 TGCTCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1785 GCTGATGACATCTCCCTGCTGAAGTGAAGAGGCCCATGGGAGAGATAGAGATTCGCCCT 1844
Db 1441 GCTGATGACATCTCCCTGCTGAAGTGAAGAGGCCCATGGGAGAGATAGAGATTCGCCCT 1500
QY 1845 GGACCAACCTCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1904
Db 1501 GGACCAACCTCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
QY 1905 CAGAGCACTCAGGACCTCCGCCACCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1964
Db 1561 CAGAGCACTCAGGACCTCCGCCACCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1965 GGCTGCAAGTGGGTTCCAGGAGTGTACCTGTAGGAACAGAGAAAGAGAAAGAG 2024
Db 1621 GGCTGCAAGTGGGTTCCAGGAGTGTACCTGTAGGAACAGAGAAAGAGAAAGAG 1680
QY 2025 CACTCTGCTGGGGAATACTTCTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTGCT 2084
Db 1681 CACTCTGCTGGGGAATACTTCTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTGCT 1740
QY 2085 TGAACCTTACGCCCCGTGAACCTTTGCTCCACCATTCCTTTTAAATTCCTCAACCCCAAGTATT 2144
Db 1741 TGAACCTTACGCCCCGTGAACCTTTGCTCCACCATTCCTTTTAAATTCCTCAACCCCAAGTATT 1800

QY 2145 CTTCTTTTCTTTAGTTTTCAGAAGTACTGGCATCACACGACAGTTACCTTTGGCGTGTCTCCC 2204
Db 1801 CTTCTTTTCTTTAGTTTTCAGAAGTACTGGCATCACACGACAGTTACCTTTGGCGTGTCTCCC 1860
QY 2205 TGTGTTACCTTGGCAGAGAGAGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAAGTACGA 2264
Db 1861 TGTGTTACCTTGGCAGAGAGAGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCAAGTACGA 1920
QY 2265 GAGGATGCACAGCTTTGCTATTTGCTTTAGAGACAGGAGCTGTATTAACAAGCCTAACATT 2324
Db 1921 GAGGATGCACAGCTTTGCTATTTGCTTTAGAGACAGGAGCTGTATTAACAAGCCTAACATT 1980
QY 2325 GGTGCAAAAGATTGCCTCTTTGAATT 2348
Db 1981 GGTGCAAAAGATTGCCTCTTTGAATT 2004

RESULT 10

US-09-845-226-1
: Sequence 1, Application US/09845226
: Patent No. US20020115600A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Jordan J.N.
: APPLICANT: Hong, Lin
: APPLICANT: Ghosh, Arun K.
: TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
: FILE REFERENCE: OMRF 182
: CURRENT APPLICATION NUMBER: US/09/845,226
: CURRENT FILING DATE: 2001-04-30
: PRIOR APPLICATION NUMBER: 09/603,713
: PRIOR FILING DATE: 2000-06-27
: PRIOR APPLICATION NUMBER: 60/168,060
: PRIOR FILING DATE: 1999-11-30
: PRIOR APPLICATION NUMBER: 60/177,836
: PRIOR FILING DATE: 2000-01-25
: PRIOR APPLICATION NUMBER: 60/178,368
: PRIOR FILING DATE: 2000-01-27
: PRIOR APPLICATION NUMBER: 60/210,292
: PRIOR FILING DATE: 2000-06-08
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 3252
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-845-226-1

Query Match 85.3%; Score 2004; DB 10; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 GCGGAGTGTGCTGCTGCGCCACGCGCACCCAGACGGATCCGGCTGCCCTCGCGCAGCGC 404
Db 1 GCGGAGTGTGCTGCTGCGCCACGCGCACCCAGACGGATCCGGCTGCCCTCGCGCAGCGC 60
QY 405 CTGGGGGCGCCCCCTTGGGGCTGCGGCTGCCCGGGAGACCCAGACGAGAGAGAGAGAGAG 464
Db 61 CTGGGGGCGCCCCCTTGGGGCTGCGGCTGCCCGGGAGACCCAGACGAGAGAGAGAGAGAG 120
QY 465 CCGGGCGGAGGGGAGGAGCTTTGTGGAGATGGTGGACAACTGAGGGGCAAGTCGGGGGAG 524
Db 121 CCGGGCGGAGGGGAGGAGCTTTGTGGAGATGGTGGACAACTGAGGGGCAAGTCGGGGGAG 180
QY 525 GGCTACTACGTGGAGATGACCGTGGCAGCCCCCGCAGAGCTCAACATCCTGCTGCAT 584
Db 181 GGCTACTACGTGGAGATGACCGTGGCAGCCCCCGCAGAGCTCAACATCCTGCTGCAT 240
QY 585 ACAGGAGCAGTAACCTTTGCTGAGTGGTGTGCCCGCCACCCCTTCCCTGCTGCTGCTAC 644
Db 241 ACAGGAGCAGTAACCTTTGCTGAGTGGTGTGCCCGCCACCCCTTCCCTGCTGCTGCTAC 300
QY 645 CAGAGGAGCTGTCCAGCAGACATACCGGGAGCTCCGGAGGGGTGTGTATGTGCCCTACACC 704

Db 301 CAGAGGAGCTGTCAGACACATACCGGAGCCTCCGGAAAGGTTGATGTGCCCTACACC 360
Oy 705 CAGGCAAGTGGGAAGGAGCTGGGACCGACCTGGTAAGCATCCCATGGCCCCAAC 764
Db 361 CAGGCAAGTGGGAAGGAGCTGGGACCGACCTGGTAAGCATCCCATGGCCCCAAC 420
Oy 765 GTCACTGTGCGTGGCAACATCTGTCGCATCACTGAATCAGACAAGTCTTTCATCAACGGC 824
Db 421 GTCACTGTGCGTGGCAACATCTGTCGCATCACTGAATCAGACAAGTCTTTCATCAACGGC 480
Oy 825 TCCAACTGGGAAGGATCCTGGGGTGGCCTATGCTGAGATTGCCAGGCTGACACTCC 884
Db 481 TCCAACTGGGAAGGATCCTGGGGTGGCCTATGCTGAGATTGCCAGGCTGACACTCC 540
Oy 885 CTGGAGCCTTTCTTTGACTCTCTGTAAAGCAGACCCACGTTCCCAACCTCTTTCCTCCCTG 944
Db 541 CTGGAGCCTTTCTTTGACTCTCTGTAAAGCAGACCCACGTTCCCAACCTCTTTCCTCCCTG 600
Oy 945 CAGCTTTCTGTGTGCTGCTTCCCTCAACCACTGCTGAAGTCTGGCCTCTGTGCGAGGG 1004
Db 601 CAGCTTTCTGTGTGCTGCTTCCCTCAACCACTGCTGAAGTCTGGCCTCTGTGCGAGGG 660
Oy 1005 AGCATGATCATTTGAGGATGACACCACTCGCTGTACAGGAGTCTCTGTGATACACC 1064
Db 661 AGCATGATCATTTGAGGATGACACCACTCGCTGTACAGGAGTCTCTGTGATACACC 720
Oy 1065 ATCCGGGGGAGTGTATTTAGAGTGATCATTTGCGGGTGGAGATCAATGGACAGGAT 1124
Db 721 ATCCGGGGGAGTGTATTTAGAGTGATCATTTGCGGGTGGAGATCAATGGACAGGAT 780
Oy 1125 CTGAAATGGACTGCAAGGAGTACAACTATGACAAAGACATTTGTGACAGTGGCACACC 1184
Db 781 CTGAAATGGACTGCAAGGAGTACAACTATGACAAAGACATTTGTGACAGTGGCACACC 840
Oy 1185 AACCTTCCTTTGCCCAAGAAAGTGTGAAAGCTGACGTCAAAATCCATCAAGGACGCTCC 1244
Db 841 AACCTTCCTTTGCCCAAGAAAGTGTGAAAGCTGACGTCAAAATCCATCAAGGACGCTCC 900
Oy 1245 TCCAGGAGAGTTCCTTCATGTTCTGCTAGGAGACGCTGCTGCTGGCAAGCA 1304
Db 901 TCCAGGAGAGTTCCTTCATGTTCTGCTAGGAGACGCTGCTGCTGGCAAGCA 960
Oy 1305 GGCACACCCCTTGAACATTTTCCAGTCACTCTACCTACTACCTAATGGGTAGGTTACC 1364
Db 961 GGCACACCCCTTGAACATTTTCCAGTCACTCTACCTACTACCTAATGGGTAGGTTACC 1020
Oy 1365 AACAGTCTTCCGATCACCATCCTTCCGAGCAATACCTGCGGCCAGTGGAAATGTG 1424
Db 1021 AACAGTCTTCCGATCACCATCCTTCCGAGCAATACCTGCGGCCAGTGGAAATGTG 1080
Oy 1425 GCCAGTCCCAAGAGGACTGTTACAAAGTTTGCATCTCACAGTCACTCCAGGGCACTGTT 1484
Db 1081 GCCAGTCCCAAGAGGACTGTTACAAAGTTTGCATCTCACAGTCACTCCAGGGCACTGTT 1140
Oy 1485 ATGGAGCTGTATCATGGAGGCTTCTACGTTGCTTTGATCGGGCCCGAAACGAATT 1544
Db 1141 ATGGAGCTGTATCATGGAGGCTTCTACGTTGCTTTGATCGGGCCCGAAACGAATT 1200
Oy 1545 GCGTTTGTCTCAGGCTTGGCATGTGCAAGATGAGTTTCCAGGACGCGCGTGGAAAGC 1604
Db 1201 GCGTTTGTCTCAGGCTTGGCATGTGCAAGATGAGTTTCCAGGACGCGCGTGGAAAGC 1260
Oy 1605 CCTTTTGTACCTTGGACATGGAAGACTGTGCTACACATCTCCACAGACATGAGTCA 1664
Db 1261 CCTTTTGTACCTTGGACATGGAAGACTGTGCTACACATCTCCACAGACATGAGTCA 1320
Oy 1665 ACCCTCATGACATAGCCTATGCTATGCTGCCATCTGCGCCCTCTTCTATGCTGCCACTC 1724
Db 1321 ACCCTCATGACATAGCCTATGCTATGCTGCCATCTGCGCCCTCTTCTATGCTGCCACTC 1380
Oy 1725 TGCCTCATGGTGTGTCAGTGGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1784

Db 1381 TGCCTCATGGTGTGTCAGTGGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1440
Oy 1785 GCTGATGACATCTCCCTGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1844
Db 1441 GCTGATGACATCTCCCTGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500
Oy 1845 GGACACACCTCCGTTGCTTACCTTTGGTTCACAAAGTAGGAGACACAGATGGCACCTGTGGC 1904
Db 1501 GGACACACCTCCGTTGCTTACCTTTGGTTCACAAAGTAGGAGACACAGATGGCACCTGTGGC 1560
Oy 1905 CAGACACCTCAGGACCTCCACCCACCAATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1964
Db 1561 CAGACACCTCAGGACCTCCACCCACCAATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1620
Oy 1965 GGCTGGCAAGTGGTTCACAGGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAG 2024
Db 1621 GGCTGGCAAGTGGTTCACAGGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAG 1680
Oy 2025 CACTCTGCTGGCGGGAATACCTTGTGTCACCTCAATTTAAATTTCTCCAAACCAAGTATT 2084
Db 1681 CACTCTGCTGGCGGGAATACCTTGTGTCACCTCAATTTAAATTTCTCCAAACCAAGTATT 1740
Oy 2085 TGAACCTTCAGCCCTGAACCTTTGTCACCTCAATTTAAATTTCTCCAAACCAAGTATT 2144
Db 1741 TGAACCTTCAGCCCTGAACCTTTGTCACCTCAATTTAAATTTCTCCAAACCAAGTATT 1800
Oy 2145 CTCTCTTTCTTGTAGTTTCAAGTACTGTCATCACAGCAGTGTACCTTGGCGTGTGCTCC 2204
Db 1801 CTCTCTTTCTTGTAGTTTCAAGTACTGTCATCACAGCAGTGTACCTTGGCGTGTGCTCC 1860
Oy 2205 TGTGTACCTCTGGCAGAGAACAGCAAGCTTTGTTTCCCTGCTGGCCAAAGTCACTAGGA 2264
Db 1861 TGTGTACCTCTGGCAGAGAACAGCAAGCTTTGTTTCCCTGCTGGCCAAAGTCACTAGGA 1920
Oy 2265 GAGGATGACAGTTCCTTATTTGCTTTAGAGACAGGACTGTATTAACAGCCTTAACATT 2324
Db 1921 GAGGATGACAGTTCCTTATTTGCTTTAGAGACAGGACTGTATTAACAGCCTTAACATT 1980
Oy 2325 GGTGCAAGATTGCTCTTGAATT 2348
Db 1981 GGTGCAAGATTGCTCTTGAATT 2004

RESULT 11

US-09-794-927-5
; Sequence 5, Application US/09794927
; Patent No.. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794, 927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1977


```
QY 1386 ATCTTCCGAGCAATACCTCGCGCCAGTGAAGATGTGCCACGTCGCCAAGACGACTGT 1445
Db 1006 ATCTTCCGAGCAATACCTCGCGCCAGTGAAGATGTGCCACGTCGCCAAGACGACTGT 1065
QY 1446 TACAAGTTTCCCATCTCACAGTATCATCCACGGGCACCTGTTATGGAGCTGTATCATGGAG 1505
Db 1066 TACAAGTTTCCCATCTCACAGTATCATCCACGGGCACCTGTTATGGAGCTGTATCATGGAG 1125
QY 1506 GCGTCTACCTTGTCTTTGATCGGGCCGAAACAGAAATGGCTTTGCTGTGTCAGCGCTTGC 1565
Db 1126 GCGTCTACCTTGTCTTTGATCGGGCCGAAACAGAAATGGCTTTGCTGTGTCAGCGCTTGC 1185
QY 1566 CATGTGCACGATGAGTTACGAGCGGACGGGTGGAAGCCCTTTTGTACCTTTGGACATG 1625
Db 1186 CATGTGCACGATGAGTTACGAGCGGACGGGTGGAAGCCCTTTTGTACCTTTGGACATG 1245
QY 1626 GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCTCTCATGACCATAGCCAT 1685
Db 1246 GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCTCTCATGACCATAGCCAT 1305
QY 1686 GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG 1745
Db 1306 GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCCTCATGGTGTGTCAGTGG 1365
QY 1746 CGCTGCTCCGCTCGCTCGCCAGCAGCATGATGACTTTTCTGATGACATCTCCCTGCTG 1805
Db 1366 CGCTGCTCCGCTCGCTCGCCAGCAGCATGATGACTTTTCTGATGACATCTCCCTGCTG 1425
QY 1806 AAGTGAGGAGGCCATGCGGAGAGATAGAGATTCCTCTGACACACACTCCGTTGGTTCA 1865
Db 1426 AAGTGAGGAGGCCATGCGGAGAGATAGAGATTCCTCTGACACACACTCCGTTGGTTCA 1485
QY 1866 CTTTGTGTCAAGTATGAGAGACAGATGGCACCTGTGGCCAGAGCAGCCTCAGAGCCCTCC 1925
Db 1486 CTTTGTGTCAAGTATGAGAGACAGATGGCACCTGTGGCCAGAGCAGCCTCAGAGCCCTCC 1545
QY 1926 CCACCCACCAAAATGCCCTCTGCTGTTAGTGAAGAAAGGCTGGCAAGGTGGTTCCAG 1985
Db 1546 CCACCCACCAAAATGCCCTCTGCTGTTAGTGAAGAAAGGCTGGCAAGGTGGTTCCAG 1605
QY 1986 GGACTGTACCTGTAGGAAACAGAAAGAGAGAAAGAGCAGTCTGCTGGCGGGAATACT 2045
Db 1606 GGACTGTACCTGTAGGAAACAGAAAGAGAGAAAGAGCAGTCTGCTGGCGGGAATACT 1665
QY 2046 CTTGGTCACCTCAAAATTAAGTCGGGAAATTCGTCTCTTGAACCTTTCAGCCCTGAACTT 2105
Db 1666 CTTGGTCACCTCAAAATTAAGTCGGGAAATTCGTCTCTTGAACCTTTCAGCCCTGAACTT 1725
QY 2106 TTGTCCACCAATTCCTTTAAATTCCTCAACCCAAAGATTTCTTTTCTTTAGTTTCAGAA 2165
Db 1726 TTGTCCACCAATTCCTTTAAATTCCTCAACCCAAAGATTTCTTTTCTTTAGTTTCAGAA 1785
QY 2166 GTACTGGCATCACAGCAGGTACCTTTGGCGTGTGTCCTCTGTTGGTACCTTGGCAGAGAAG 2225
Db 1786 GTACTGGCATCACAGCAGGTACCTTTGGCGTGTGTCCTCTGTTGGTACCTTGGCAGAGAAG 1845
QY 2226 AGACCAAGCTTGTTCCTCTGTCGCAAAAGTCACTAGAGAGGATGCACAGTTTGTCTATT 2285
Db 1846 AGACCAAGCTTGTTCCTCTGTCGCAAAAGTCACTAGAGAGGATGCACAGTTTGTCTATT 1905
QY 2286 TGCTTTAGACAGGGGACTGTATAAACAAGCCTTAACATTTGGTGCAGAAATTTGCCCTTTGA 2345
Db 1906 TGCTTTAGACAGGGGACTGTATAAACAAGCCTTAACATTTGGTGCAGAAATTTGCCCTTTGA 1965
QY 2346 A 2346
Db 1966 A 1966
```

RESULT 14

US-09-794-748-5

; Sequence 5, Application US/0979478

; Patent No. US20020037315A1

```
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280JL
; CURRENT APPLICATION NUMBER: US/09/794, 748
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416, 901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155, 493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404, 133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101, 594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-794-748-5
```

Query Match 80.1%; Score 1881; DB 10; Length 1977;

Best Local Similarity 96.3%; Pred. No. 0;

Matches 1966; Conservative 0; Mismatches 0; Indels 75; Gaps 1;

```
QY 306 ATGGCCCAAGCCCTGCCCTGCTGCTGTGGATGGCGGGAGGTGCTGCCCTGCCAC 365
Db 1 ATGGCCCAAGCCCTGCCCTGCTGCTGTGGATGGCGGGAGGTGCTGCCCTGCCAC 60
QY 366 GGCACCCAGACGCGCATCCGCTGCCCTTGGCAGCGCCCTGGGGGGCGCCCTGGGG 425
Db 61 GGCACCCAGACGCGCATCCGCTGCCCTTGGCAGCGCCCTGGGGGGCGCCCTGGGG 120
QY 426 CTGGGCTGCCCGGGAGACCGAGAGCCCGAGAGCCCGAGAGCCCGCGGAGGCGAGCTT 485
Db 121 CTGGGCTGCCCGGGAGACCGAGAGCCCGAGAGCCCGCGGAGGCGAGCTT 180
QY 486 GTGGAGATGTGGACAACCTGAGGGGCAAGTGGGGGAGGCTACTAGCTGGAGATGACC 545
Db 181 GTGGAGATGTGGACAACCTGAGGGGCAAGTGGGGGAGGCTACTAGCTGGAGATGACC 240
QY 546 GTGGGAGCCCCCGCAGACGCTCAACATCCTGTGGATACAGCAGCAGTAACCTTGCA 605
Db 241 GTGGGAGCCCCCGCAGACGCTCAACATCCTGTGGATACAGCAGCAGTAACCTTGCA 300
QY 606 GTGGGTGCTGCCCGCCCGCCCTTCTTCATCGCTACTACACAGAGCAGCTGTCCAGCACA 665
Db 301 GTGGGTGCTGCCCGCCCGCCCTTCTTCATCGCTACTACACAGAGCAGCTGTCCAGCACA 360
QY 666 TACCGGACCTCCGGAGGGTGTGTATGTGCCCTTACACCCAGGCAAGTGGGAGGGGAG 725
Db 361 TACCGGACCTCCGGAGGGTGTGTATGTGCCCTTACACCCAGGCAAGTGGGAGGGGAG 420
QY 726 CTGGGACCGACCTGGTAAGCATCCCATGCCCCCAACGTCAGCTGCGTGCCCAACATT 785
Db 421 CTGGGACCGACCTGGTAAGCATCCCATGCCCCCAACGTCAGCTGCGTGCCCAACATT 480
QY 786 GCTGCCATCACTGAATCAGACAAGTTCTTCAATCAACGGCTCCAACCTGGGAAGGATCTCTG 845
Db 481 GCTGCCATCACTGAATCAGACAAGTTCTTCAATCAACGGCTCCAACCTGGGAAGGATCTCTG 540
QY 846 GGGCTGGCTATGCTGAGATGGCCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 905
Db 541 GGGCTGGCTATGCTGAGATGGCCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 566
```


Qy	906	CTGGTAAAGCAGACACCCAGTTC	CCCAACCTTCTTCTCCCTGCAGCTTTGTGGTGTGCTGCCTTC	905
Db	567	-----	-----GCTTTGTGGTGGCTTC	585
Qy	966	CCCCTAACCAAGTCTGAAGTGTGGCTCTGTCTGGAGGAGCATGATCATTTGGAGTATC		1025
Db	586	CCCCTAACCAAGTCTGAAGTGTGGCTCTGTCTGGAGGAGCATGATCATTTGGAGTATC		645
Qy	1026	GACCACTCGCTGTATACACAGGCAGTCTCTGGTATACACCCATCCGGCGGAGTGGTATTAT		1085
Db	646	GACCACTCGCTGTATACACAGGCAGTCTCTGGTATACACCCATCCGGCGGAGTGGTATTAT		705
Qy	1086	GAGGTGATCATTTGCGGGTGGAGATCAATGACACAGGATCTGAAATGGACTGCAAGGAG		1145
Db	706	GAGGTGATCATTTGCGGGTGGAGATCAATGACACAGGATCTGAAATGGACTGCAAGGAG		765
Qy	1146	TACAACTATGACAACGAGCATCTGTGACAGTGGCACACCAACCTTCGTTTGCCCAAGAA		1205
Db	766	TACAACTATGACAACGAGCATCTGTGACAGTGGCACACCAACCTTCGTTTGCCCAAGAA		825
Qy	1206	GTGTTTGAAGCTGCAGTCAAAATCCAAATCCAAAGCAGCCTCTCCACGGAGAAAGTTCCCTGAT		1265
Db	826	GTGTTTGAAGCTGCAGTCAAAATCCAAATCCAAAGCAGCCTCTCCACGGAGAAAGTTCCCTGAT		885
Qy	1266	GGTTCTTGCTTAGGAGACAGCTGGTGTGCTGGCAAGCAGGACACACCCCTTGGACATTT		1325
Db	886	GGTTCTTGCTTAGGAGACAGCTGGTGTGCTGGCAAGCAGGACACACCCCTTGGACATTT		945
Qy	1326	TTCCCAAGTCATCTCACTACTACCTAATGGTGAGGTTTACCAACAGTCTTCGCACTCAC		1385
Db	946	TTCCCAAGTCATCTCACTACTACCTAATGGTGAGGTTTACCAACAGTCTTCGCACTCAC		1005
Qy	1386	ATCTCTCCGACAGCAATACCTCGGCCAGTGGAAAGATGTGGCCACGTCCTCAAGACGACTGT		1445
Db	1006	ATCTCTCCGACAGCAATACCTCGGCCAGTGGAAAGATGTGGCCACGTCCTCAAGACGACTGT		1065
Qy	1446	TACAAGTTTGCCATCTCACAGTCATCCACGGGCATGTTTATGGGAGCTGTTATCATCGGAG		1505
Db	1066	TACAAGTTTGCCATCTCACAGTCATCCACGGGCATGTTTATGGGAGCTGTTATCATCGGAG		1125
Qy	1506	GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAAACGAATTGGCTTTGCTCTCAGCGCTTGC		1565
Db	1126	GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAAACGAATTGGCTTTGCTCTCAGCGCTTGC		1185
Qy	1566	CATGTGCACGATGAGTTCCAGGACGCGGGTGGAAAGGCCCTTTTGTCACTTTGGACATG		1625
Db	1186	CATGTGCACGATGAGTTCCAGGACGCGGGTGGAAAGGCCCTTTTGTCACTTTGGACATG		1245
Qy	1626	GAAGACTGTGCTACAAATTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT		1685
Db	1246	GAAGACTGTGCTACAAATTTCCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT		1305
Qy	1686	GTCATGGCTGCCATCTGGCGCCTCTTTCATGTCGCCACTCTGCTCATGGTGTGTCAAGTGG		1745
Db	1306	GTCATGGCTGCCATCTGGCGCCTCTTTCATGTCGCCACTCTGCTCATGGTGTGTCAAGTGG		1365
Qy	1746	CGTGCTTCCGCTGGCTCGGCCACAGCATGATGACTTTGTGTGATGACATCTCCCTGTGTG		1805
Db	1366	CGTGCTTCCGCTGGCTCGGCCACAGCATGATGACTTTGTGTGATGACATCTCCCTGTGTG		1425
Qy	1806	AAGTGAGGAGGCCATGGCAGAGATAGAGATTCCTCTGGACACACCTTCGTTGTTCA		1865
Db	1426	AAGTGAGGAGGCCATGGCAGAGATAGAGATTCCTCTGGACACACCTTCGTTGTTCA		1485
Qy	1866	CTTTTGGTCACAAGTAGGACACAGATGGCACCTCTGGCCAGGACACCTTCAGGACCCCTCC		1925
Db	1486	CTTTTGGTCACAAGTAGGACACAGATGGCACCTCTGGCCAGGACACCTTCAGGACCCCTCC		1545
Qy	1926	CCACCCACCAATGCTCTGCTTGTATGGAGAAGAAAGGCTGGCAAGGTGGGTTCACAG		1985
Db	1546	CCACCCACCAATGCTCTGCTTGTATGGAGAAGAAAGGCTGGCAAGGTGGGTTCACAG		1605

RESULT 15

RESULT IS
US-09-794-925-5

03-03-794-323-3
; Sequence 5, Application US/097949225

; sequence 3, application 03;
; Patent No. US20020064819A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

APPLICANT: Bienkowski, Michael J.

APPLICANT: HEINRIKSON, Robert L.

; APPLICANT: Parodi, Luis A.

APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S

; TITLE OF INVENTION: THEREFOR

FILE REFERENCE: 28341/6280HI

; CURRENT APPLICATION NUMBER: US/09/07860A
; CURRENT FILING DATE: 2001-02-27

; CURRENT FILING DATE: 2001-02-27
 ; PRIOR APPLICATION NUMBER: 00/415 801

;; PRIOR APPLICATION NUMBER: 09/410
: PRIOR FILING DATE: 1998-10-13

; PRIOR FILING DATE: 1999-10-13
 ; PRIOR APPLICATION NUMBER: 60/155 193

;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23

; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR APPLICATION NUMBER: 09/404,133
 ; PRIOR FILING DATE: 1999-09-23

; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881

PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594

PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 1977

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-794-925-5

Query Match	Score 1881;	DB 10;	Length 1977;
Best local similarity	0.628	Prod No. 0.	

306	QY	ATGGCCCAAGCCCTGCCCTCGCTTCCTGCTGTGATGGCGCGGGGAGTGTCTGCTGTCCAC	365
	Db		
1		ATGGCCCAAGCCCTGCCCTCGCTTCCTGCTGTGATGGCGCGGGGAGTGTCTGCTGTCCAC	60
366	QY	GGCACCCAGCAGCGGATCCGGTGGCCCTCGCCAGCGGCCCTGGGGGGGCGCCCGCTGGGG	425
61	db	GGCACCCAGCAGCGGATCCGGTGGCCCTCGCCAGCGGCCCTGGGGGGGCGCCCGCTGGGG	120

Qy	426	CTCGGCTGCCCGGAGACCCGACGAAGAGCCCGGAGAGCCCGCGCGAGGGGACGCTTT	485
Db	121	CTCGGCTGCCCGGAGACCCGACGAAGAGCCCGGAGAGCCCGCGAGGGGACGCTTT	180
Qy	486	GTGGAGATGTTGGACAACTGAGGGCAAGTCGGGGCAGGCGTACTACGTGGACATCACC	545
Db	181	GTGGAGATGTTGGACAACTGAGGGCAAGTCGGGGCAGGCGTACTACGTGGACATCACC	240
Qy	546	GTGGGAGCCCCCGCAGACGCTCAACATCTCTGTGGATACAGGCAGCAGTAACCTTTGCA	605
Db	241	GTGGGAGCCCCCGCAGACGCTCAACATCTCTGTGGATACAGGCAGCAGTAACCTTTGCA	300
Qy	606	GTGGTGTCTGCCCCCACCCTTCTCTGCTACCTACTACAGAGCAGCTGTCAGCACACA	665
Db	301	GTGGTGTCTGCCCCCACCCTTCTCTGCTACCTACTACAGAGCAGCTGTCAGCACACA	360
Qy	666	TACCGGACCTCCGGAAGGCTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG	725
Db	361	TACCGGACCTCCGGAAGGCTGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG	420
Qy	726	CTGGGCACCGACCTGGTAAGCATCCCCATGSCCCCAACGTCACGTGCGTGCCCAACATT	785
Db	421	CTGGGCACCGACCTGGTAAGCATCCCCATGSCCCCAACGTCACGTGCGTGCCCAACATT	480
Qy	786	GCTGCCATCACTGAATCAGACAGTTCTTCATCAGGCTCCAACTGGGAAGGCACTCGT	845
Db	481	GCTGCCATCACTGAATCAGACAGTTCTTCATCAGGCTCCAACTGGGAAGGCACTCGT	540
Qy	846	GGCTGGCTTATGCTCAGATTCGAGGCTCGACACTCCCTGGAGCCTTCTTTTGACTCT	905
Db	541	GGCTGGCTTATGCTCAGATTCGAGGCTCCAG-----	566
Qy	906	CTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGGCTTC	965
Db	567	-----GCCTTGTGGTGGCTGGCTTC	585
Qy	966	CCCTCAACGAGTCTCAAGTGTGGCCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC	1025
Db	586	CCCTCAACGAGTCTCAAGTGTGGCCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC	645
Qy	1026	GACCACTCGCTGTACACAGGCACTCTGGTATACACCCATCCGCGGGAGTGGTATTAT	1085
Db	646	GACCACTCGCTGTACACAGGCACTCTGGTATACACCCATCCGCGGGAGTGGTATTAT	705
Qy	1086	GAGTGATCATTTGCGGGTGGAGATCAATGACAGGATCTGAAATGGACTGCAAGGAG	1145
Db	706	GAGTGATCATTTGCGGGTGGAGATCAATGACAGGATCTGAAATGGACTGCAAGGAG	765
Qy	1146	TACAACATACAAAGAGCATTTGGACAGTGGCACACCACTTCGTTTCCCAAGAAA	1205
Db	766	TACAACATACAAAGAGCATTTGGACAGTGGCACACCACTTCGTTTCCCAAGAAA	825
Qy	1206	GTGTTTGAAGCTCAGTCAAATCCATCAAGGCAGCCTCCTCCAGGAAAGTTCCCTGAT	1265
Db	826	GTGTTTGAAGCTCAGTCAAATCCATCAAGGCAGCCTCCTCCAGGAAAGTTCCCTGAT	885
Qy	1266	GTTTTCTGGCTAGGAGAGCAGTGTGTGCTGGCAAGCAGGCACACCCCTTTGGAACATT	1325
Db	886	GGTTTCTGGCTAGGAGAGCAGTGTGTGCTGGCAAGCAGGCACACCCCTTTGGAACATT	945
Qy	1326	TTCCCACTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCACTCCTTCGGCATCACC	1385
Db	946	TTCCCACTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCACTCCTTCGGCATCACC	1005
Qy	1386	ATCCTTTCCGACGAATACCTTGGCGGCAGTGGAAAGATGTGGCACCTCCCAAGACGACTGT	1445
Db	1006	ATCCTTTCCGACGAATACCTTGGCGGCAGTGGAAAGATGTGGCACCTCCCAAGACGACTGT	1065
Qy	1446	TACAAGTTTCCCATCTCACAGTCATCCACGGGCATGTTTATGGAGCTGTTTATCATGGAG	1505
Db	1066	TACAAGTTTCCCATCTCACAGTCATCCACGGGCATGTTTATGGAGCTGTTTATCATGGAG	1125

QY	1506	GGCTTCTACGTTGTCTTTGATCGGGCCCGAAACGAATTGGCTTTGCTGTCAAGCGCTTGC	1505
Db	1126	GGCTTCTACGTTGTCTTTGATCGGGCCCGAAACGAATTGGCTTTGCTGTCAAGCGCTTGC	1185
QY	1566	CATGTGCACGATGAGTTCAGGACGCGCAGCGGTGGAAGGCCCTTTTGTCAACCTTGGACATG	1625
Db	1186	CATGTGCACGATGAGTTCAGGACGCGCAGCGGTGGAAGGCCCTTTTGTCAACCTTGGACATG	1245
QY	1626	GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT	1685
Db	1246	GAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT	1305
QY	1686	GTCATGCTGCATCTGCGGCCCTCTTTCATGCTGCCACTCTGCTCTCATGCTGTGTCAAGTGG	1745
Db	1306	GTCATGCTGCATCTGCGGCCCTCTTTCATGCTGCCACTCTGCTCTCATGCTGTGTCAAGTGG	1365
QY	1746	CGCTGCCCTCCGCTGCCCTGCCACGACGATGATGACTTTTGTGATGACATCTCCCTGCTGTG	1805
Db	1366	CGCTGCCCTCCGCTGCCCTGCCACGACGATGATGACTTTTGTGATGACATCTCCCTGCTGTG	1425
QY	1806	AAGTGAGGAGGCCCATGGGCAGAAAGATACAGATTCCCCTGGACACACCTCGGTGGTTTCA	1865
Db	1426	AAGTGAGGAGGCCCATGGGCAGAAAGATACAGATTCCCCTGGACACACCTCGGTGGTTTCA	1485
QY	1866	CTTTTGGTTCACAAGTAGGAGACACAGATGGCACTGTGGCCAGACGACCTCAGGACCCCTCC	1925
Db	1486	CTTTTGGTTCACAAGTAGGAGACACAGATGGCACTGTGGCCAGACGACCTCAGGACCCCTCC	1545
QY	1926	CCACCCACCAATGCCCTCTGCCCTTGATGAGAAAGGCTGGCAAGGTGGGTTCAG	1985
Db	1546	CCACCCACCAATGCCCTCTGCCCTTGATGAGAAAGGCTGGCAAGGTGGGTTCAG	1605
QY	1986	GGACTGTACCTGTAGGAACAGAAAGAGAAAGCAAGCACTCTGCTGGCGGGAATACT	2045
Db	1606	GGACTGTACCTGTAGGAACAGAAAGAGAAAGCAAGCACTCTGCTGGCGGGAATACT	1665
QY	2046	CTTGGTCACCTCAAAATTAAGTCGGGAAATCTGCTGTCTGAAACTTCAGCCCTGAACT	2105
Db	1666	CTTGGTCACCTCAAAATTAAGTCGGGAAATCTGCTGTCTGAAACTTCAGCCCTGAACT	1725
QY	2106	TTGTGTCCACCATCTCTTAAATCTCCACCCAAAGTATTCTTCTTTCTTAGTTTCAGAA	2165
Db	1726	TTGTGTCCACCATCTCTTAAATCTCCACCCAAAGTATTCTTCTTTCTTAGTTTCAGAA	1785
QY	2166	GTACTGGCATCACACGAGGTTTACCTTGGCGTGTGCCCTGTGGTACCTGCGCAGAGAAG	2225
Db	1786	GTACTGGCATCACACGAGGTTTACCTTGGCGTGTGCCCTGTGGTACCTGCGCAGAGAAG	1845
QY	2226	AGACCAAGCTTGTTCCTGCTGGCCAAAGTCAAGTAGGAGAGAGTGCACAGTTTGTCTATT	2285
Db	1846	AGACCAAGCTTGTTCCTGCTGGCCAAAGTCAAGTAGGAGAGAGTGCACAGTTTGTCTATT	1905
QY	2286	TGCTTTTAGAGACAGGGACTCTATAAACAGCCCTAACATTTGGTGCAAGATTTGCCCTTTGA	2345
Db	1906	TGCTTTTAGAGACAGGGACTCTATAAACAGCCCTAACATTTGGTGCAAGATTTGCCCTTTGA	1965
QY	2346	A 2346	
Db	1966	I 1966	

Search completed: March 2, 2003, 13:26:27
Job time : 123 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2003, 14:26:30 ; Search time 2167.5 Seconds

(without alignments)
17544.177 Million cell updates/sec

Title: US-09-723-722a-44

Perfect score: 2348

Sequence: 1 ccacgcgcgcctcacagc.....caaaattgcctctgaatt 2348

Scoring table: IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1649	70.2	3859	11 AK014464	AK014464 Mus muscu
2	990	42.2	1114	9 AF150387	AF150387 AF150387
3	954.2	40.6	2503	11 AK014390	AK014390 Mus muscu
4	819.6	34.9	913	14 BQ949536	BQ949536 AGENCOURT
5	740.6	31.5	1059	13 BI084083	BI084083 602869445
6	728	31.0	747	14 BQ018588	BQ018588 UT-H-DH1

C	7	723.8	30.8	865	13	BI084878	BI084878 602869445
C	8	713.8	30.4	761	13	BM016288	BM016288 603642760
C	9	688.6	29.3	700	9	AL570757	AL570757 AL570757
C	10	688.6	29.3	904	10	BE378929	BE378929 601237528
C	11	681.4	29.0	718	12	BF057767	BF057767 7K53e11.x
C	12	664.4	28.3	945	14	BQ669630	BQ669630 AGENCOURT
C	13	657.8	28.0	684	12	BF338951	BF338951 602036021
C	14	647.4	27.6	730	13	BM048418	BM048418 603625683
C	15	634.2	27.0	840	12	BE885725	BE885725 601508937
C	16	631.4	26.9	799	12	BF204338	BF204338 601866411
C	17	624	26.6	635	10	BE296749	BE296749 601175134
C	18	598.6	25.5	684	10	AV725513	AV725513 AV725513
C	19	595	25.3	623	13	BI828125	BI828125 603075385
C	20	594.8	25.3	862	13	BI157614	BI157614 602922587
C	21	579	24.7	720	12	BG288435	BG288435 602383404
C	22	574.2	24.5	649	12	BE837592	BE837592 RC2-PN009
C	23	563.8	24.0	644	12	BE783981	BE783981 601472451
C	24	556.4	23.7	939	12	BF203806	BF203806 601868788
C	25	533	22.7	836	9	AL544727	AL544727 AL544727
C	26	524	22.3	525	9	AI005033	AI005033 ou91b12.x
C	27	514.4	21.9	541	10	AW770546	AW770546 h186c06.x
C	28	511.4	21.8	513	14	BQ637035	BQ637035 he04g12.y
C	29	508	21.6	864	12	BE872035	BE872035 601448124
C	30	476.2	20.3	605	12	BE793449	BE793449 601588510
C	31	475.2	20.2	514	12	BF439471	BF439471 nab65a03
C	32	472	20.1	487	13	BM313609	BM313609 ig72h03.y
C	33	464	19.8	464	9	AA701598	AA701598 zi34h07.s
C	34	464	19.8	464	9	AI127789	AI127789 qc32c02.x
C	35	453.2	19.3	458	9	AL700814	AL700814 DKFZp686H
C	36	451.2	19.2	542	10	AV667139	AV667139 AV667139
C	37	447.4	19.1	537	13	BG927600	BG927600 HNC43-1-E
C	38	445	19.0	461	9	AL700831	AL700831 DKFZp686I
C	39	445	19.0	493	12	BF443234	BF443234 260654 MA
C	40	444	18.9	444	9	AI095556	AI095556 qb19f04.x
C	41	443	18.9	443	9	AI127822	AI127822 qc36f02.x
C	42	426.4	18.2	428	9	AI094243	AI094243 qa43c11.s
C	43	425.6	18.1	1092	13	BM454007	BM454007 AGENCOURT
C	44	424.2	18.1	573	12	BE799078	BE799078 601588381
C	45	423.6	18.0	447	10	AW015855	AW015855 UI-H-BIO-

ALIGNMENTS

RESULT 1
AK014464
LOCUS
DEFINITION
AK014464 3859 bp mRNA linear HTC 19-JAN-2002
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:4122401C04:beta-site APP cleaving enzyme, full insert sequence.
ACCESSION
AK014464
VERSION
AK014464.1 GI:12852334
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain:C57BL/6J) 16 days embryo head cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library
clone:4122401C04.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1
Carninci,P. and Hayashizaki,Y.

TITLE
High-efficiency full-length cDNA cloning

JOURNAL
Meth. Enzymol. 303, 19-44 (1999)

MEDLINE
99279253

PUBMED
10349636

REFERENCE

2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE
20499374

PUBMED
11042159


```
QY 479 CAGCTTTGGAGATGTTGACAACTGAGGGCAAGTCGGGCGAGGCTACTACGTGGA 538
Db 603 CAGCTTTGGAGATGTTGACAACTGAGGGCAAGTCGGGCGAGGCTACTATGTGA 662
QY 539 GATGACCTGGGAGAGCCCGGAGAGCGCTCAACATCCTGGTGGATACAGGCAGCAATA 598
Db 663 GATGACCTGAGGAGAGCCCGGAGAGCGCTCAACATCCTGGTGGAGACAGGCGAGTAGTA 722
QY 599 CTTTGGAGTGGGTGCTCCCGCCCAACCTTCCCTGATCGTACTACACAGGAGCGCTGC 658
Db 723 CTTTGGAGTGGGTGCTCCCGCCCAACCTTCCCTGATCGTACTACACAGGAGCGCTGC 782
QY 659 CAGCACATACCGGAGACTCCCGAAGGTGTATGTGCCCTTACACCCAGGCGAAGTGGGA 718
Db 783 CAGCACATACGAGACCTCCCGAAGGTGTATGTGCCCTTACACCCAGGCGAAGTGGGA 842
QY 719 AGGGAGCTGGGACCGACCTGTTAAGCATCCCGCATCGGCGCCCAACGTCTACTGTGCGTGC 778
Db 843 GGGGGAACCTGGGACCGACCTGTTAAGCATCCCGCATCGGCGCCCAACGTCTACTGTGCGTGC 902
QY 779 CAACATTGCTGCCATCACTGAATCAGACAGTCTTCTCATCAACGGCTCCAACCTGGGAAGG 838
Db 903 CAACATTGCTGCCATCACTGAATCAGACAGTCTTCTCATCAACGGTCTCCAACCTGGGAAGG 962
QY 839 CATCTGGGCTGGCTTATGCTGAGATTGCCAGGCTGAGACCTCCCTGGAGCCCTTTCTT 898
Db 963 CATCTAGGCTGGCTTATGCTGAGATTGCCAGGCTGAGACCTCCCTGGAGCCCTTTCTT 1022
QY 899 TGACTCTCTGTTAAGCAGACCCAGCTTCCGAACTCTTCTCCCTGAGCTTTGTGTGTC 959
Db 1023 TGACTCTCTGTTAAGCAGACCCAGCTTCCGAACTCTTCTCCCTGAGCTTTGTGTGTC 1082
QY 959 TGGCTTCCCTCAACCAAGCTGTAAGTGTGGCTCTGCGAGGGAGCATGATCATTTGG 1018
Db 1083 TGGCTTCCCTCAACCAAGCTGTAAGTGTGGCTCTGCGAGGGAGCATGATCATTTGG 1142
QY 1019 AGTATCGACCACTCGCTGTACACAGCAGTCTCTGGTATACACCCATCCGCGGGAGTG 1078
Db 1143 TGGTATCGACCACTCGCTGTACACAGCAGTCTCTGGTATACACCCATCCGCGGGAGTG 1202
QY 1079 GTATTATGAGTGTATGTCGGGTGGAGATCAATGGACAGATCTGAAATGGAGTG 1138
Db 1203 GTATTATGAGTGTATGTCGGGTGGAGATCAATGGTCAAGATCTCAAGATGGAGTG 1262
QY 1139 CAAGAGTACAACTATGACAAGAGCATTGTGGACAGTGGACACCAACCTTCGTTTGGC 1198
Db 1263 CAAGAGTACAACTATGACAAGAGCATTGTGGACAGTGGACACCAACCTTCGTTTGGC 1322
QY 1199 CAAGAAAGTGTGGAAGTGTGCAATCCATCAAGGAGCCTCTCCAGGAGAGATG 1258
Db 1323 CAAGAAAGTGTGGAAGTGTGCAATCCATCAAGGAGCCTCTCCAGGAGAGATG 1382
QY 1259 CCCTGATGGTTCGTGGCTAGGAGAGCAGTGTGCTGTGGCAAGCAGGACCCCTTG 1318
Db 1383 CCCGGATGGCTTTGGCTAGGAGAGCAGTGTGCTGTGGCAAGCAGGACCCCTTG 1442
QY 1319 GAACATTTTCCAGTCTACTCTACCTAATGGGTGAGTTTACCAACCAAGTCTCTCCG 1378
Db 1443 GAACATTTTCCAGTCTACTCTACCTAATGGGTGAGTTTACCAACCAAGTCTCTCCG 1502
QY 1379 CATCACCATCTTCCGAGCAATACCTGCGGCCAGTGGAGATGTGGCCAGTCCCAAGA 1438
Db 1503 CATCACCATCTTCTCAGCAATACCTACGCGCGTGGAGACCTGGCCAGTCCCAAGA 1562
QY 1439 CGACTGTTACAACTTGGCATCTCACAGTCAATCCACGGGCACTGTTATGGAGCTGTAT 1498
Db 1563 CGACTGTTACAACTTGGCATCTCACAGTCAATCCACGGGCACTGTTATGGAGCGCTAT 1622
QY 1499 CATGAGGCTTCTACGTTGTCTTGTGTCGGGCCGCAAGCAATGTGCTTTGTGTCAG 1558
Db 1623 CATGAGGCTTCTATGTCGTTCTGATCGAGCCCGAAGCAATGTGCTTTGTGTCAG 1682
QY 1559 CGCTTGCCATGTGACGATGAGTTTCAGGACGCGAGCGGTGGAAGGCCCTTTTGTACCTT 1618
```

```
Db 1683 CGCTTGCATGTGCACCATGAGTTTCAGAGCGGCGAGTGAAGGTCGTTGTACGGC 1742
QY 1619 GGACATGGAGAGCTGTGGCTACACATTTCCACAGACAGATGAGTCAACCCTCATGACAT 1678
Db 1743 AGACATGGAGAGCTGTGGCTACACATTTCCACAGACAGATGAGTCAACACTATGACCAT 1802
QY 1679 AGCTATGTCTATGGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCCCTCATGGTGTG 1738
Db 1803 AGCTATGTCTATGGCGGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCCCTCATGGTGTG 1862
QY 1739 TCAGTGGCGTGGCTCCGCTGCCCTGCCAGCAGCATGATGATGATGATGATGATGATGATGATG 1798
Db 1863 TCAGTGGCGTGGCTGCCCTGCCCTGCCAGCAGCATGATGATGATGATGATGATGATGATGATG 1922
QY 1799 CTTGCTCAAGTGAAGGAGGCCATGGGCGAGATAGAGATTCCTCTGGACCACTCCG 1858
Db 1923 CTTGCTCAAGTGAAGGAGGCCATGGGCGAGATAGAGATTCCTCTGGACCACTCCG 1982
QY 1859 TGGTTCACTTTGGTGCACA--AGTAGGAGACACAGATGGCAGCTGTGGCCAGACCACTCA 1916
Db 1983 TGGTTCCCTTTGGTGCACATGAGTTGGAGCTATGATGATGATGATGATGATGATGATGATG 2042
QY 1917 GGACCTTCCCAACCCCAAAATGCTCTGCCCTTGGATGGAGAGGAAAGGCTGCAAGGT 1976
Db 2043 GGACCTTCCCAACCCCAAAATGCTCTGCCCTTGGATGGAGAGGAAAGGCTGCAAGGT 2101
QY 1977 GGGTTCCAGGAGCTGTACCTGTAGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2036
Db 2102 GGGTTCCAGGAGCTGTACCTGTAGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2161
QY 2037 GGGAAATCTCTTGGTGCACCTCAAAATTTAAGTCGGGAAATCTGCTGTGAACCTTCAGC 2096
Db 2162 AGGAATATCTTGGTGCACCTCAAAATTTAAGTCGGGAAATCTGCTGTGAACCTTCAGC 2220
QY 2097 CCTGAACCTTTGTCCCACTTCTTAAATTCCTCAACCCCAAGATATCTTCTTTCTTA 2156
Db 2221 CCTGACCTCTGCCCCAGCGCTTTAGAGTCTCCAACCTAAAGTATTTCTTTATGTC--- 2277
QY 2157 GTTTCAGAGTACTGGCATCACAGGTTACCTTGGCGTGTCTCCCTGTGGTGTGACCTG 2216
Db 2278 CTTCCAGAGTACTGGCATCACAGGTTACCTTGGCGTGTCTCCCTGTGGTGTGACCTG 2336
QY 2217 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2276
Db 2337 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2393
QY 2277 TTTGCTATTTGCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2336
Db 2394 TTTGCTATTTGCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2452
QY 2337 GCCTCTTGA 2345
Db 2453 GCGTCTTGA 2461
```

RESULT 2

AF150387

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AF150387 1114 bp mRNA linear EST 22-JUN-1999
AF150387 Human mRNA from cd34+ stem cells Homo sapiens cdna clone
CBMAPH03, mRNA sequence.
AF150387
AF150387.1 GI:5133823
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1114)
Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L.,
Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.
Human mRNA from cd34+ stem cells
Unpublished (1999)

COMMENT

Contact: Ye M
Shanghai Institute of Hematology
Shanghai Second Medical University, Rui-Jin Hospital
197 Rui-Jin Road II, Shanghai, 200025, P. R. China
Email: zchen@sh.cn

FEATURES
source

Location/Qualifiers
1. .ll14
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CBMAPH03"
/clone_lib="Human mRNA from cd34+ stem cells"
/tissue_type="umbilical cord blood"
/cell_type="cd34+ stem cell"
/note="cloned by differential display method after
chemical induction of terminal differentiation of cell
line"

BASE COUNT 275 a 300 c 286 g 253 t

ORIGIN

Query Match 42.2%; Score 990; DB 9; Length 1114;
Best Local Similarity 97.3%; Pred. No. 4.8e-232;
Matches 1092; Conservative 0; Mismatches 20; Indels 10; Gaps 8;

QY 1017 GGAGGTATCGACCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCGGCGGGAG 1076
DB 1 GGAGGTATCGACCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCGGCGGGAG 60
QY 1077 TGGTATTATGAGGTGATCATTTCTGCGGGTGGAGATCAATGACAGGATCTCAAAATGGAC 1136
DB 61 TGGTATTATGAGGTGATCATTTCTGCGGGTGGAGATCAATGACAGGATCTCAAAATGGAC 120
QY 1137 TGCAAGGAGTACAACTATGACAAAGAGCATTTGGACAGTGGACACCAACCTTCGTTTG 1196
DB 121 TGCAAGGAGTACAACTATGACAAAGAGCATTTGGACAGTGGACACCAACCTTCGTTTG 180
QY 1197 CCCAAGAAAGTGTGAGAGTCCAGTCAATCCATCAAGCAGGCTCTCTCCACGGAGAG 1256
DB 181 CCCAAGAAAGTGTGAGAGTCCAGTCAATCCATCAAGCAGGCTCTCTCCACGGAGAG 240
QY 1257 TTCCCTGATGTTTCTGGCTAGGAGAGCAGTGGTGTCTGGCAAGCAGGACCAACCCCT 1316
DB 241 TTCCCTGATGTTTCTGGCTAGGAGAGCAGTGGTGTCTGGCAAGCAGGACCAACCCCT 300
QY 1317 TGGAACTTTTCCAGTCAATCTACTCTACCTAATGGGTGAGGTACCAACAGTCTTC 1376
DB 301 TGGAACTTTTCCAGTCAATCTACTCTACCTAATGGGTGAGGTACCAACAGTCTTC 360
QY 1377 CGCATCACCACTCTCCGAGCAATACCTCCGCCAGTGGAGATGGCCACGTCCTCCAA 1436
DB 361 CGCATCACCACTCTCCGAGCAATACCTCCGCCAGTGGAGATGGCCACGTCCTCCAA 419
QY 1437 GACGACTGTTTACAGTTTGGCAATCTCACAGTATCCACGGGCACTGTTATGGGAGCTGT 1496
DB 420 GACGACTGTTTACAGTTTGGCAATCTCACAGTATCCACGGGCACTGTTATGGGAGCTGT 478
QY 1497 ATCATGGAGGCTTCTA-CGTTGTCTTTGATCGGGCCCGCAAAAGCAATGGCTTTGCTGT 1555
DB 479 ATCATGGAGGCTTCTA-CGTTGTCTTTGATCGGGCCCGCAAAAGCAATGGCTTTGCTGT 538
QY 1556 CAGCGCTTGCCATGTGCAGGATGAGTTTCAGGAGCGGCGGTGGAAGCCCTTTTGTCTAC 1615
DB 539 CAGCGCTTGCCATGTGCAGGATGAGTTTCAGGAGCGGCGGTGGAAGCCCTTTTGTCTAC 597
QY 1616 CTTGCACATGGAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGAC 1675
DB 598 CTTGCACATGGAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGAC 657
QY 1676 CATAGCCTATGTGATGGTGGCCATCTGCGCCCTCTTCATGCTGCCACTTCGCTCATGGT 1735
DB 658 CATAGCCTATGTGATGGTGGCCATCTGCGC-CCTCTTCATGCTGCCACTTCGCTCATGGT 716
QY 1736 GTGTCAG-TGGGCTGCCTCCCTGCCTGGCCAGCAGCATGATGACTTTCCTGATGACA 1794
DB 1736 GTGTCAG-TGGGCTGCCTCCCTGCCTGGCCAGCAGCATGATGACTTTCCTGATGACA 1794

DB 717 GTGTACAGATGGCGCTGCCCTCCGCTGCTGTGGCCAGCAGCATGATGACTATGCTGATGACA 776
QY 1795 TCTCCTCTGCTGAAGTGAAGGAGGCGCCATGGCAGAGATAGAGATTCCCTCTGCACACACC 1854
DB 777 TCTCCTCTGCTGAAGTGAAGGAGGCGCCATGGCAGAGATAGAGATTCCCTCTGCACACACC 836
QY 1855 TCCGTGTGTTTCACTTTGGTTCACAAAGTAGGAGACACAGATGGCACCCTGTGCCAGAGACCT 1914
DB 837 TACGTGGTCACTTTGGCAC--AGTAGGACACAGATGGCACCCTGTGCCAGAGACCT 893
QY 1915 CAGGACCTCTCCACCCACCAATGCTCTGCTTGTATGGAGAGAGAAAGCTGGCAAG 1974
DB 894 CAGGACCTCTCCACCCACCAATGCTCTGCTTGTATGGAGAGAGAAAGCTGGCAAG 953
QY 1975 GTGGTTCAGGAGTGTACCTGTAGGAAACAGAGAAAGAGAGAAAGAGCACTCTGCTG 2034
DB 954 GTGGTTCAGGAGTGTACCTGTAGGAAACAGAGAAAGAGAGAAAGAGCACTCTGCTG 1012
QY 2035 CGGGGAATACCTTTGGTCACTTCAAAATTTAAGTCGGGAAATTTCTGCTCTTGAACCTTCA 2094
DB 1013 CGGGGAATACCTTTGGTCACTTCAAAATTTAAGTCGGGAAATTTCTGCTCTTGAACCTTCA 1072
QY 2095 GCCTGGAACCTTTGTGCCACCATTCCTTTAAATTTCTTCAACCC 2136
DB 1073 ACCCTGAACCTTTGTCAACCATTCCTTTAAATTTATACAACCC 1114
RESULT 3
AK014390 2503 bp mRNA linear HTc 19-JAN-2002
LOCUS Mus musculus adult male brain cDNA, RIKEN full-length enriched
DEFINITION library, clone:3526402A15:beta-site APP cleaving enzyme, full
insert sequence.
ACCESSION AK014390
VERSION AK014390.1 GI:12852207
KEYWORDS HTc; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male brain cDNA to mRNA,
clone.lib:RIKEN full-length enriched mouse cDNA library
ORGANISM clone:3526402A15.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,


```

Db 721 CCCACCAATGCCCTCGCTTGTGGAGAGGAAAAAGCTGGCAGGTGNGTTCCAGGG 780
Qy 1988 ACTGTACTGTAGGAAACAGAGAAAGAGAGAAAGACACTCTGCTGGCGGGAATATCTC- 2046
Db 781 ACTGTACCTGTAGAACNCCNAGAGAGAGAAAGACACTCTGCTGGCGGGAATATCTCN 840
Qy 2047 TTGTGTCACCTCAAAATTTAAGTCGGGAATTTCT-GCTGCTTGAACACTTTCAGCCC 2098
Db 841 TTGTGTCACCTCAAAATTTAAGTCGGGAATTTCTGCTGCTTGAACACTTTCAGGCC 893

RESULT 5
BI084083 1059 bp mRNA linear EST 20-JUN-2001
LOCUS 602869445F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013994 5',
DEFINITION mRNA sequence.
ACCESSION BI084083
VERSION BI084083.1 GI:14502413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1059)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1820 row: j column: 11
High quality sequence stop: 794.
FEATURES
Location/Qualifiers
1..1059
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5013994"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGCAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 243 a 300 c 273 g 243 t
ORIGIN
Query Match 31.5%; Score 740.6; DB 13; Length 1059;
Best Local Similarity 92.5%; Pred. No. 7.5e-171;
Matches 837; Conservative 0; Mismatches 54; Indels 14; Gaps 5;
Qy 1157 CAAGAGCATTTGGACAGTGGCACCACCAACCTTCGTTTGGCCCAAGAAAGTGTGTAAGC 1216
Db 2 CAAGAGCATTTGGACAGTGGCACCACCAACCTTCGTTTGGCCCAAGAAAGTGTGTAAGC 61
Qy 1217 TGCAGTCAAAATCCATCAAGGAGGAGCTCTCCACGAGAGAGTTCCCTGTATGTTTGTGGCT 1276
Db 62 TGCAGTCAAAATCCATCAAGGAGGAGCTCTCCACGAGAGAGTTCCCTGTATGTTTGTGGCT 121
Qy 1277 AGGAGAGCAGTGTGTGCTGGCAGGAGGAGCACCACCTTTGGACATTTTCCAGTGCAT 1336
Db 122 AGGAGAGCAGTGTGTGCTGGCAGGAGGAGCACCACCTTTGGACATTTTCCAGTGCAT 181

```

```

Qy 1337 CTCACTCTACCTAATGGGTGAGGTTACCAACAGTCCCTTCCGCATCACCATCTTCCGCA 1396
Db 182 CTCACTCTACCTAATGGGTGAGGTTACCAACAGTCCCTTCCGCATCACCATCTTCCGCA 241
Qy 1397 GCATACCTTGGCGGCGAGTGGAGATGTGGCCACGTCCTCCCAAGACGACTGTTTACAAGTTTGC 1456
Db 242 GCAATACCTTGGCGGCGAGTGGAGATGTGGCCACGTCCTCCCAAGACGACTGTTTACAAGTTTGC 301
Qy 1457 CATCTCACAGTCATCCACGGGCACTGTTTATGGAGCTCTTATCATGGAGGCTTCTAGCT 1516
Db 302 CATCTCACAGTCATCCACGGGCACTGTTTATGGAGCTTATCATGGAGGCTTCTAGCT 361
Qy 1517 TGTCTTTGATCGGCGCCGAAACGAATTTGCTGCTCAGCGCTTCCCATGTGCACGA 1576
Db 362 TGTCTTTGATCGGCGCCGAAACGAATTTGCTGCTCAGCGCTTCCCATGTGCACGA 421
Qy 1577 TGAGTTTCAGGACGGCAGCGGTGGAGGCCCTTTTGTCACTTTGGACATGGAGACTGTGG 1636
Db 422 TGAGTTTCAGGACGGCAGCGGTGGAGGCCCTTTTGTCACTTTGGACATGGAGACTGTGG 481
Qy 1637 CTACAACATTCCACAGACAGATGAGTCAACCTCATGACCATAGCTATGTCACTGGCTGC 1696
Db 482 CTACAACATTCCACAGACAGATGAGTCAACCTCATGACCATAGCTATGTCACTGGCTGC 541
Qy 1697 CATCTGCGCCCTCTTCATGCTGCCACTCTGCTCATGTGTGTCACTGGCGCTGCTCCG 1756
Db 542 CATCTGCGCCCTCTTCATGCTGCCACTCTGCTCATGTGTGTCACTGGCGCTGCTCCG 601
Qy 1757 CTGCTTGGCGCAGCAGCATGATGACTTTGCTGTAGTACATCTCCCTGTGTAAGTGAGGAGG 1816
Db 602 CTGCTTGGCGCAGCAGCATGATGACTTTGCTGTAGTACATCTCCCTGTGTAAGTGAGGAGG 661
Qy 1817 CCCATGGCACAAGATAGAGATTCCTCTGGACACACCTCCGTTGTTTGTGTACACA 1876
Db 662 CCCATGGCACAAGATAGAGATTCCTCTGGACACACCTCCGTTGTTTGTGTACACA 721
Qy 1877 AGTAGGAGACAGATGGCAGCTGTGGCAGAGCAGCAGCCTCCACCCACCACCA 1936
Db 722 AGTAGGAGACAGATGGCAGCTGTGGCAGAGCAGCAGCCTCCACCAACCA 778
Qy 1937 ATGCTCTGCTTGTATGAGAGAAAGAGGCTGGCAAGTGGTTCACGAGGACTTACCT 1996
Db 779 ATGCTCTGTG-CTTGTATGAGAGCCG-ACAGGCTGGCAGCAGTGGGTTCC---GGACTGTACCT 833
Qy 1997 GTAGAAACAGAAAAAGAGAAAGAACACTCTGCTGCGGGAATCTCTTGTGTACCT 2056
Db 834 GTAGAAACAGCA-----AGCAAGAGAAAGAGTCTCTGTGCGGGAATCTCTTGTGTACCTC 887
Qy 2057 CAAAT 2061
Db 888 AATTT 892

RESULT 6
BI018588/c 747 bp mRNA linear EST 27-MAR-2002
LOCUS UI-H-DH1-awu-c-12-0-UI-s1 NCI_CGAP_DH1 Homo sapiens cDNA clone
DEFINITION IMAGE:5823683 3', mRNA sequence.
ACCESSION BI018588
VERSION BI018588.1 GI:19753865
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 747)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende

```

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 1-22, >AT-rich<low_complexity
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Source

Location/Qualifiers

1..747

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5823683"

/clone_lib="NCI_CGAP_DH1"

/tissue_type="Metastatic Chondrosarcoma"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I;

NCI_CGAP_DH1 is a normalized cDNA library containing the

following tissue(s): VS-8 Cell line from Metastatic

Chondrosarcoma in Lung. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT73-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is AGATCATTCG.

TAG_LIB=UI-H-DH1

TAG_TISSUE=Lung

TAG_SEQ=AGATCATTCG

BASE COUNT 176 a 178 c 192 g 200 t 1 others

ORIGIN

Query Match

Best Local Similarity 31.0%; Score 728; DB 14; Length 747;

Matches 728; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1620 GACATGAAGACTGTGGCTACACATTCACAGACAGATGAGTCAACCCCTCATGACCATTA 1679
|||||
DB 747 GACATGAAGACTGTGGCTACACATTCACAGACAGATGAGTCAACCCCTCATGACCATTA 688
|||||
QY 1680 GCCTATGTGTCATGCGCTGCGCCCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1739
|||||
DB 687 GCNTATGTGTCATGCGCTGCGCCCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
|||||
QY 1740 CAGTGGCGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCC 1799
|||||
DB 627 CAGTGGCGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCC 568
|||||
QY 1800 CTGCTGAAGTGAAGGCGCCATGGCGAGAAGATAGAGATTCCTCCCTGACACACCTCCGT 1859
|||||
DB 567 CTGCTGAAGTGAAGGCGCCATGGCGAGAAGATAGAGATTCCTCCCTGACACACCTCCGT 508
|||||
QY 1860 GGTTCACCTTTGTCACAACTAGGACACACAGATGGCACCTGTGGCCAGACGACCTCAGGA 1919
|||||
DB 507 GGTTCACCTTTGTCACAACTAGGACACACAGATGGCACCTGTGGCCAGACGACCTCAGGA 448
|||||
QY 1920 CCCTCCCAACCCACCAATGCTCTGCTCTGATGGAGAAGAAAGCTGGCAAGGTGGG 1979
|||||
DB 447 CCCTCCCAACCCACCAATGCTCTGCTCTGATGGAGAAGAAAGCTGGCAAGGTGGG 388
|||||
QY 1980 TTCCAGGACTGTACCTCTAGGAACACAGAAAAGAGAAAGAGCACTCTGCTGGCGGG 2039
|||||
DB 387 TTCCAGGACTGTACCTCTAGGAACACAGAAAAGAGAAAGAGCACTCTGCTGGCGGG 328
|||||
QY 2040 AATACTCTTGCTACCTCAATTTAAGTCGGGAAATTCGCTCTGCTCAAACTTCAGCCCT 2099
|||||
DB 327 AATACTCTTGCTACCTCAATTTAAGTCGGGAAATTCGCTCTGCTGCTGCTGCTGCTGCTGCT 268
|||||

QY 2100 GAACCTTTGTCCACCATTCTTTAAATCTTCCAAACCAAGATATCTCTTTCTTTAGTT 2159
|||||
DB 267 GAACCTTTGTCCACCATTCTTTAAATCTTCCAAACCAAGATATCTCTTTCTTTAGTT 208
|||||
QY 2160 TCAGAACTACTGGCATCACACGAGGTACCTTTGGCGTGTGCCCTGTGGTACCTCGCA 2219
|||||
DB 207 TCAGAACTACTGGCATCACACGAGGTACCTTTGGCGTGTGCCCTGTGGTACCTCGCA 148
|||||
QY 2220 GAGAAGAGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCACTAGGAGGAGTGCACACTTT 2279
|||||
DB 147 GAGAAGAGACCAAGCTTTGTTCCCTGCTGGCCAAAGTCACTAGGAGGAGTGCACACTTT 88
|||||
QY 2280 GCTATTTGCTTTAGACAGACGAGGACTGTATAAACAAGCCCTTAACATTTGCTGCAAGATTTGCC 2339
|||||
DB 87 GCTATTTGCTTTAGACAGACGAGGACTGTATAAACAAGCCCTTAACATTTGCTGCAAGATTTGCC 28
|||||
QY 2340 TCTTGAATT 2348
|||||
DB 27 TCTTGAATT 19
|||||
RESULT 7
BI084878/c
LOCUS
DEFINITION
602859445T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013994 3',
mRNA sequence.
ACCESSION
BI084878
VERSION
BI084878.1 GI:14503208
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 865)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI820 row: j column: 11
High quality sequence start: 18
High quality sequence stop: 821.
Location/Qualifiers
1..865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5013994"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 207 a 224 c 228 g 205 t 1 others
ORIGIN
Query Match 30.8%; Score 723.8; DB 13; Length 865;
Best Local Similarity 95.3%; Pred. No. 9.7e-167;
Matches 810; Conservative 0; Mismatches 33; Indels 7; Gaps 6;

Qy	1413	GTGGAAGATGTCGGCCAGCTGCCAAGACGACTGTTACAAAGTTTGCCATCTCACATCATCC	1472
Db	865	GTGGAAGATGTCGGCCAGCTGCCAAGACGACTGTTAC-AGGTTGCCATCTCACATCATCC	807
Qy	1473	ACGGGCACTGTTATGGAG-CTGTTATATCGAGGGCTTCTACGTTGTCTTTGATCGGGC	1531
Db	806	ACGGCAACTGTTATGGAGGCTGTTATCATATGGGAGGCTTCTACGTTGTCTTTGATCGGGC	747
Qy	1532	CCGAAAACGAATGGCTTTGCTGTCAGCGCTTGCCA-TGTGCAGATGAGTTCAGGACGG	1590
Db	746	CCGAAAACGAATGGCTTTGCTGTCAGCGCTTGCCATGTGCACAGATGAGTTCAGGACGG	687
Qy	1591	CAGCGTGGAGGCCCTTTTGTACCTTGACATGGAAGACTGTGGC--TACAACATTC	1648
Db	686	CAGCGTGGAGGCCCTTTTGTCACTTGGGACATGGAAGACTGTGGCATACANACATTC	627
Qy	1649	ACAGACAGATGAGTCAACCCCTCATGACCATAGCCATATGTATGCTGCGCATCTCGGCCCT	1708
Db	626	ACAGACAGATGAGTCAACCCCTCATGACCATAGCCATATGTATGCTGCGCATCTCGGCCCT	567
Qy	1709	CTTCATCTGCACTCTGCTGTCAGCTGCTGTCAGTGGCGCTGCCCTCGCTGCGGCCA	1768
Db	566	CTTCATCTGCACTCTGCTGTCAGCTGCTGTCAGTGGCGCTGCCCTCGCTGCGGCCA	507
Qy	1769	GCAGCATGATGACTTTGCTGATGACATCTCCCTGCTGAAAGTGAGGAGGCCCATGGCAGA	1828
Db	506	GCAGCATGATGACTTTGCTGATGACATCTCCCTGCTGAAAGTGAGGAGGCCCATGGCAGA	447
Qy	1829	AGATAGAGATTCCTCGTGGACACACCTCCGTTGTTTCACTTTGGTCAACAAGTAGGAGAC	1888
Db	446	AGATAGAGATTCCTCGTGGACACACCTCCGTTGTTTCACTTTGGTCAACAAGTAGGAGAC	388
Qy	1889	AGATGGCACCTGTGGCCAGACACCTCAGGACCCCTCCACCACCAAAATGCTCTGCCT	1948
Db	387	AGATGGCACCTGTGGCCAGACACCTCAGGACCCCTCCACCACCAAAATGCTCTGCCT	328
Qy	1949	TGATGGAGAAGAAAGGCTGGCAAGTGGTTCCAGGGACTGTACCTGTAGGAACAGA	2008
Db	327	TGATGGAGAAGAAAGGCTGGCAAGTGGTTCCAGGGACTGTACCTGTAGGAACAGA	268
Qy	2009	AAAGAGAAGAAAGACACTGCTGTGGGGAATACTCTTGGTCACTCAAAATTTAAGTC	2068
Db	267	AAAGAGAAGAAAGACACTGCTGTGGGGAATACTCTTGGTCACTCAAAATTTAAGTC	208
Qy	2069	GGGAAATTCGCTCTTGAACCTTTCAGCCCTGAACCTTTGTCACCATTCCTTTAAATTC	2128
Db	207	GGGAAATTCGCTCTTGAACCTTTCAGCCCTGAACCTTTGTCACCATTCCTTTAAATTC	148
Qy	2129	TCCAACCCAAAGTATCTCTTTTCTTAGTTTCAGAAAGTACTGGCATCACACGAGTTA	2188
Db	147	TCCAACCCAAAGTATCTCTTTTCTTAGTTTCAGAAAGTACTGGCATCACACGAGTTA	88
Qy	2189	CTTTGGGCTGTGCTCCCTGTGTACCCCTGGCAGAGAAGACCAAGCTGTTTCCCTGCTG	2248
Db	87	CTTTGGGCTGTGCTCCCTGTGTGTA-CTTGGCAGAGAAGACCAAGCTGTTTCCCTGCTG	29
Qy	2249	GCCAAAGTCA 2258	
Db	28	GCCAAAGTCA 19	
RESULT 8			
BM016288			
LOCUS	603642760F1 NIH_MGC_87 Homo sapiens	cdna clone IMAGE:5418687 5',	761 bp mRNA linear EST 30-OCT-2001
DEFINITION	mRNA sequence.		
ACCESSION	BM016288		
VERSION	BM016288.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

REFERENCE	1 (bases 1 to 761)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-r@mail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12068 row: h column: 16 High quality sequence stop: 728.		
FEATURES	Location/Qualifiers		
source	1..761		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5418687"		
	/clone_lib="NIH_MGC_87"		
	/tissue_type="mammary adenocarcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		
BASE COUNT	170 a	221 c	197 g 171 t 2 others
ORIGIN	Query Match 30.4%; Score 713.8; DB 13; Length 761; Best Local Similarity 99.3%; Pred. No. 2.7e-164; Matches 726; Conservative 0; Mismatches 4; Indels 1; Gaps 1;		
Qy	1221	GTCAAAATCCATCAAGGACGCTCCTCCACGGAGAAGTTCCCTGATGTTCTTGGCTAGGA	1280
Db	1	GTCAAAATCCATCAAGGACGCTCCTCCACGGAGAAGTTCCCTGATGTTCTTGGCTAGGA	60
Qy	1281	GAGCAGCTGGTGTGCTGGCAAGCAGGACACCCCTTGGAAACATTTCCCACTCATCTCA	1340
Db	61	GAGCAGCTGGTGTGCTGGCAAGCAGGACACCCCTTGGAAACATTTCCCACTCATCTCA	120
Qy	1341	CTCTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCATCACCATCTTCCCGCAGCA	1400
Db	121	CTCTACCTTAATGGGTGAGGTTACCAACCACTCTTCCGCATCACCATCTTCCCGCAGCA	180
Qy	1401	TACCTGGCGGCAGTGGAAAGATGTGGCCACGCTCCCAAGACGACTGTTACAAGTTTCCCATC	1460
Db	181	TACCTGGCGGCAGTGGAAAGATGTGGCCACGCTCCCAAGACGACTGTTACAAGTTTCCCATC	240
Qy	1461	TCACAGTATCCACGGGCACCTGTTATGGAGCTGTATATCGAGGCTTCTACGTTGTC	1520
Db	241	TCACAGTATCCACGGGCACCTGTTATGGAGCTGTATATCGAGGCTTCTACGTTGTC	300
Qy	1521	TTTGTATCGGCGCCGAAAACGAAATTTGGCTTTGCTGTCAGCGCTTGGCATGTGCACCATGAG	1580
Db	301	TTTGTATCGGCGCCGAAAACGAAATTTGGCTTTGCTGTCAGCGCTTGGCATGTGCACCATGAG	360
Qy	1581	TTCAGGACGGCAGCGGTGGAAAGGCCCTTTTGTACCTTTGGACATGGAAGACTGTGGCTAC	1640
Db	361	TTCAGGACGGCAGCGGTGGAAAGGCCCTTTTGTACCTTTGGACATGGAAGACTGTGGCTAC	420
Qy	1641	AACATTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCTATGTATGCTGCCATC	1700
Db	421	AACATTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCTATGTATGCTGCCATC	480
Qy	1701	TGGCGCCTCTTCATGCTGCCACTCTGCCTCATGTGTGTGTCAGTGGCGCTTGCCTCGCTGC	1760
Db	481	TGGCGCCTCTTCATGCTGCCACTCTGCCTCATGTGTGTGTCAGTGGCGCTTGCCTCGCTGC	540
Qy	1761	CTGGCGCCAGCAGCATGATGACTTTTGTCTGATGACATCTCCCTGCTCAAGTGGAGGCCCCA	1820

```
Db 541 CTCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTGAGTGAGTGAGGAGGCCCA 600
QY 1821 TGGCAGAGATAGAGATTCCCTCGACACACACCTCCGCTGGTTTCACCTTTGGTTCACAAGTA 1880
Db 501 TGGCAGAGATAGAGATTCCCTCGACACACACCTCCGCTGGTTTCACCTTTGGTTCACAAGTA 660
QY 1881 GGAGACACAGATGGCACCCTGTGGCCAGAGACACCTCAGGACCCCTCCCAACCAACCAATGC 1940
Db 661 GGAGACACAGATGGCACCCTGTGGCCAGAGACACCTCAGGACCCCTCCCA-CCACAAATGC 719
QY 1941 CTCTGCCTTGA 1951
Db 720 CTCTGCCTTGA 730

RESULT 9
AL570757/c
LOCUS
DEFINITION
AL570757 LTI_NFL006_PL2 700 bp mRNA linear EST 16-FEB-2001
prime, mRNA sequence.
ACCESSION
VERSION
AL570757.1 GI:12927378
KEYWORDS
SOURCE
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 700)
Li,W.B., Gruber,C., Jessee,J. and Pollayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI022VF22"
/clone_lib="LIL_NFL006_PL2"
/tissue_type="placenta"
/notes="vector: pCMVSPORT 6; site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 169 a 173 c 191 g 161 t 6 others
ORIGIN

Query Match 29.3%; Score 688.6; DB 9; Length 700;
Best Local Similarity 99.3%; Pred. No. 4e-158;
Matches 688; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1555 TCAGCGCTTGCATGTGCACGATGATTCAGGACGCGGTGGAGGCCCTTTTGTCA 1614
Db 693 TCAGCGCTTGCATGTGCACGATGATTCAGGACGCGGTGGAGGCCCTTTTGTCA 634
QY 1615 CTTGGACATGGAAGACTGTGGCTACACATTCACAGACAGATGAGTCAACCTCATGA 1674
Db 633 CYTTGACATGGAAGACTGTGGCTACACATTCACAGACAGATGAGTCAACCTCATGA 574
QY 1675 CCATAGCCTATGTATGCTGCCCTTTCATGCTGCCACTCTGCCTCATGG 1734
Db 573 CCATAGCCTATGTATGCTGCCCTTTCATGCTGCCACTCTGCCTCATGG 514
QY 1735 TGTGTCAGTGGCGCTGCTCGCTGCCCTGCCAGCAGATGATGACTTGTGATGACA 1794
|||||
```

adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

```
BASE COUNT      191 a   240 c   256 g   217 t
ORIGIN
Query Match      29.3%; Score 688.6; DB 10; Length 904;
Best Local Similarity 98.0%; Pred. No. 4.2e-158;
Matches 740; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
QY 1054 GGTATACACCCATCCGGCGGAGTGGTATTATGAGGTG-ATCATTTGGCGGTGGAGATC 1112
      |||
Db 1 GGTATACACCCATCCGGCGGAGTGGTATTATGAGGTCCATCATTTGCGGGTGGAGATC 60
QY 1113 AATGGACAGATCTGAANAATGGAGCTGCAAGAGTACAACTATGACAAAGCATTTGTGCAC 1172
      |||
Db 61 AATGGACAGATCTGAANAATGGAGCTGCAAGAGTACAACTATGACAAAGCATTTGTGCAC 120
QY 1173 AGTGGCACCAACCAACCTTCGTTTGGCCCAAGAAAGTGTGAAAGCTGCAGTCAAAATCCATC 1232
      |||
Db 121 AGTGGCACCAACCAACCTTCGTTTGGCCCAAGAAAGTGTGAAAGCTGCAGTCAAAATCCATC 180
QY 1233 AA-GGCAGCCTCTCCACGAGAAAGTTCCCTGATGGTTTCTGGCTAGGAGACAGCTGGT 1291
      |||
Db 181 AAGGGCAGCCTCTCCACGAGAAAGTTCCCTGATGGTTTCTGGCTAGGAGACAGCTGGT 240
QY 1292 GTGCTGGCAAGCAGGACCAACCCCTTGGAAACATTTTCCAGTCACTCATCTTACCTTAAT 1351
      |||
Db 241 GTGCTGGCAAGCAGGACCAACCCCTTGGAAACATTTTCCAGTCACTCATCTTACCTTAAT 300
QY 1352 GGGTGGAGTTTACCAACAGCTCTTCCGCATCACCATCTTCCGAGCAATACCTGCGGCC 1411
      |||
Db 301 GGGTGGAGTTTACCAACAGCTCTTCCGCATCACCATCTTCCGAGCAATACCTGCGGCC 360
QY 1412 AGTGGAAAGATGTGGCCACAGCTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCAATC 1471
      |||
Db 361 AGTGGAAAGATGTGGCCACAGCTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCAATC 420
QY 1472 CACGGGCACATGTTATGGAGCTGTTATCATGGAGGGCTTCTACGTTCTCTTTGATCGGGC 1531
      |||
Db 421 CACGGGCACATGTTATGGAGCTGTTATCATGGAGGGCTTCTACGTTCTCTTTGATCGGGC 480
QY 1532 CCGAAAACGAATTGGCTTTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTTCAGGAGGCG 1591
      |||
Db 481 CCGAAAACGAATTGGCTTTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTTCAGGAGGCG 540
QY 1592 AGCGGTGGAAAGGCCCTTTTGTACCTTTGGACATGGAAGACTGTGGCTTACAACTTCCACA 1651
      |||
Db 541 AGCGGTGGAAAGGCCCTTTTGTACCTTTGGACATGGAAGACTGTGGCTTACAACTTCCACA 600
QY 1652 GACAGATGAGTCAACCCCTCATGACCATAGCCTATGTATGGCTGCCATCTGCGCCCTCTT 1711
      |||
Db 601 GACAGATGAGTCAACCCCTCATGACCATAGCCTATGTATGGCTGCCATCTGCGCCCTCTT 660
QY 1712 CATGCTGCCACTCTGCCCTCATGTGTGTGTCAGTGGCGCTGCTCGCTGCCCTGCCACGCA 1771
      |||
Db 661 CATGCTG-CACTCTGCTCATGTGTGTGTCAGTGGCGCTG---CTCCGCTGCTGCCACGCA 716
QY 1772 GCATGATGACTTTGCTGATGACATCTCTCCCTGCTGA 1806
      |||
Db 717 GCATGATGACTTTGCTGATGACATCTCTCCCTGCTGA 751
```

```
RESULT 11
BF057767/c      718 bp      mRNA      linear      EST 16-OCT-2000
LOCUS           7k53e11.x1 NCI_CGAP-G6 Homo sapiens cDNA clone IMAGE:3479156 3'
DEFINITION      similar to TR:Q9ULS1 Q9ULS1 KIAA1149 PROTEIN ;, mRNA sequence.
ACCESSION       BF057767
VERSION          BF057767.1 GI:10811663
KEYWORDS         EST.
SOURCE           human.
```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 718)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cqapbs-re@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 465.
Location/Qualifiers
1. 718
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3479156"
/clone_lib="NCI_CGAP_G6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA
from the normalized library NCI_CGAP_G64 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257086-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

```
BASE COUNT      174 a   177 c   190 g   176 t   1 others
ORIGIN
Query Match      29.0%; Score 681.4; DB 12; Length 718;
Best Local Similarity 99.3%; Pred. No. 2.4e-156;
Matches 715; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 1611 GTCACCTTGGACATGGAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAA-CCTC 1670
      |||
Db 718 GTCACCTTGGACATGGAGACTGTGGCTTAC--ACATTCACAGACAGATGAGTCAA-CCTC 661
QY 1671 ATGACCATAGCTATGTGCTGGCTGCCATCTGCGCCCTTCTTCATGCTGCCACTCTGCCTC 1730
      |||
Db 660 ATGACCATAGCTATGTGCTGGCTGCCATCTGCGCCCTTCTTCATGCTGCCACTCTGCCTC 601
QY 1731 ATGCTGTGTCAGTGGCGCTGCTCGCTGCTGCGCCAGCAGCATGATGCTTGTCTGAT 1790
      |||
Db 600 ATGCTGTGTCAGTGGCGCTGCTCGCTGCTGCGCCAGCAGCATGATGCTTGTCTGAT 541
QY 1791 GACATCTCCCTGCTGAGTGAGAGGCCCATGGGAGAGATAGAGATTCCCTCGACCA 1850
      |||
Db 540 GACATCTCCCTGCTGAGTGAGAGGCCCATGGGAGAGATAGAGATTCCCTCGACCA 481
QY 1851 CACCTCGCTGCTTCACTTTGCTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGC 1910
      |||
Db 480 CACCTCGCTGCTTCACTTTGCTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGC 421
QY 1911 ACCTCAGAGACCTCCCCACCCACCAATGCTCTGCTTGTGAGAGAGAGAAAGGCTGG 1970
      |||
Db 420 ACCTCAGAGACCTCCCCACCCACCAATGCTCTGCTTGTGAGAGAGAGAAAGGCTGG 361
QY 1971 CAAGGTGGTTCCAGGAGCTGTACTGTAGGAACAGAGAGAGAGAGAGAGAGAGCTCT 2030
      |||
Db 360 CAAGGTGGTTCCAGGAGCTGTACTGTAGGAACAGAGAGAGAGAGAGAGAGAGAGCTCT 301
```



```
QY 2031 GCTGGCGGGAATACTCTTGGTCACCTCAAAATTTAAAGTCGGGAAATTTCTGCTGTGAAAC 2090
Db 300 GCTGGCGGGAATACTCTTGGTCACCTCAAAATTTAAAGTCGGGAAATTTCTGCTGTGAAAC 241
QY 2091 TTGACCCCTGAACC-TTTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGTATTCTTCT 2149
Db 240 TTGACCCCTGAACCTTTTGTCCACCATTCCTTTAAATTTCTCAACCCAAAGTATTCTTCT 181
QY 2150 TTTCTTAGTTTCAGAAGTACTGGCATCACACGAGGTACCTTTGGGCTGTGTCCTCTGG 2209
Db 180 TTTCTTAGTTTCAGAAGTACTGGCATCACACGAGGTACCTTTGGGCTGTGTCCTCTGG 121
QY 2210 TACCCCTGGCAGAGAGACCAAGCTTTGTTCCCTGCTGCGCCAAAGTCAGTAGGAGAGGA 2269
Db 120 TACCCCTGGCAGAGAGACCAAGCTTTGTTCCCTGCTGCGCCAAAGTCAGTAGGAGAGGA 61
QY 2270 TGCACAGTTTGTCTATTGCTTTTAGAGACAGGAGCTGTATTAACAAAGCCTAACATTGTGTC 2329
Db 60 TGCACAGTTTGTCTATTGCTTTTAGAGACAGGAGCTGTATTAACAAAGCCTAACATTGTGTC 1

RESULT 12
B0669630 945 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8354446 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275949
DEFINITION 5', mRNA sequence.
ACCESSION B0669630
VERSION B0669630.1 GI:21780464
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2459 row: 0 column: 22
High quality sequence stop: 506.
FEATURES
source Location/Qualifiers
1..945
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6275949"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 253 a 231 c 245 g 214 t 2 others
ORIGIN
Query Match 28.3%; Score 664.4; DB 14; Length 945;
Best Local Similarity 99.4%; Pred. NO. 3.7e-152;
Matches 676; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1670 CATGACCATAGCCCTATGTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTTCGCT 1729
|||||
```

```
Db 1 CATGACCATAGCCCTATGTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTTCGCT 60
QY 1730 CATGGTGTGTGTCAGTGGCGCTGCCCTCGCTCGCCAGCAGCATGATGACTTTGCTGA 1789
|||||
Db 61 CATGGTGTGTGTCAGTGGCGCTGCCCTCGCTCGCCAGCAGCATGATGACTTTGCTGA 120
QY 1790 TGACATCTCCCTGCTGAAAGTGAGGAGGCCCATGGCAGAAATAGAGATTTCCCTCGACCC 1849
|||||
Db 121 TGACATCTCCCTGCTGAAAGTGAGGAGGCCCATGGCAGAAATAGAGATTTCCCTCGACCC 180
QY 1850 ACACCTCCGTTGTTTACCTTTGGTGCACAAATAGGAGACACAGATGGACCTGTGGCCAGAG 1909
Db 181 ACACCTCCGTTGTTTACCTTTGGTGCACAAATAGGAGACACAGATGGACCTGTGGCCAGAG 240
QY 1910 CACCTCAGGACCTCCGCCACCCCAATGCTCTGCTTGTATGAGAGAGAAAGGCTG 1969
|||||
Db 241 CACCTCAGGACCTCCGCCACCCCAATGCTCTGCTTGTATGAGAGAGAAAGGCTG 300
QY 1970 GCAAGGTGGGTTCCAGGAGCTGTACCTGTAGGAAACAGAAAAGAAAGAACGACCTC 2029
|||||
Db 301 GCAAGGTGGGTTCCAGGAGCTGTACCTGTAGGAAACAGAAAAGAAAGAACGACCTC 360
QY 2030 TGTGCGCGGGAATACTCTTGGTCACCTCAAAATTTAAAGTCGGGAAATTTCTGCTGTGAAA 2089
|||||
Db 361 TGTGCGCGGGAATACTCTTGGTCACCTCAAAATTTAAAGTCGGGAAATTTCTGCTGTGAAA 420
QY 2090 CTTACGCCCTGAACCTTTGTGTCACCATTTCTTAAATTTCTCCACCCCAAGTATTCTTCT 2149
|||||
Db 421 CTTACGCCCTGAACCTTTGTGTCACCATTTCTTAAATTTCTCCACCCCAAGTATTCTTCT 480
QY 2150 TTTCTTAGTTTCAGAAGTACTGGCATCACACGAGGTACCTTTGGCGTGTGTCCTCTGTTG 2209
|||||
Db 481 TTTCTTAGTTTCAGAAGTACTGGCATCACACGAGGTACCTTTGGCGTGTGTCCTCTGTTG 540
QY 2210 TACCCCT-GGCAGAGAGAGACCAAGCTTTTCCCTGCTGCGCCAAAGTCAGTAGGAGAGG 2268
|||||
Db 541 TACCCCTGGCAGAGAGAGACCAAGCTTTTCCCTGCTGCGCCAAAGTCAGTAGGAGAGG 600
QY 2269 ATGCACAGTTTGTCTATTGCTTTAGAGACAGGAGCTGTATTAACAAAGCCTAACATTGGTG 2328
Db 601 ATGCACAGTTTGTCTATTGCTTTAGAGACAGGAGCTGTATTAACAAAGCCTAACATTGGTG 660
QY 2329 CAAAGATTGCTCTTGAATT 2348
|||||
Db 661 CAAAGATTGCTCTTGAATT 680

RESULT 13
B0669630 684 bp mRNA linear EST 22-NOV-2000
LOCUS 602036021F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4184140
DEFINITION 5', mRNA sequence.
ACCESSION B0669630
VERSION B0669630.1 GI:11285371
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9501 row: 1 column: 05
```

High quality sequence stop: 652.

FEATURES

Location/Qualifiers
1. 684
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4184140"
/clone_lib="NCI_CGAP_Brn64"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/note="organ: brain; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 161 a 189 c 172 g 162 t
ORIGIN

Query Match 28.0%; Score 657.8; DB 12; Length 684;
Best Local Similarity 99.4%; Pred. No. 1.5e-150;
Matches 681; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1534 GAAACGAATTGGCTTTGGCTGTCAGCGCTTGCCATGTGCACGATGAGTTTCAGGACGGCAG 1593
DB 1 GAAACGAATTGGCTTTGGCTGTCAGCGCTTGCCATGTGCACGATGAGTTTCAGGACGGCAG 60
QY 1594 CGGTGGAAGGCCCTTTGTACCTTGGACATGGAAGACTGTGGCTACAACTTCCACAGA 1653
DB 61 CGGTGGAAGGCCCTTTGTACCTTGGACATGGAAGACTGTGGCTACAACTTCCACAGA 120
QY 1654 CAGATGAGTCAACCCCTCATGACCATGACCTATGTCATGGCTGCACCTGGCCCTCTTCA 1713
DB 121 CAGATGAGTCAACCCCTCATGACCATGACCTATGTCATGGCTGCACCTGGCCCTCTTCA 180
QY 1714 TGCTGCCACTCTGCCCTCATGCTGTGTCAGTGGCGCTGCCCTGCCCTGCCGACGAGC 1773
DB 181 TGCTGCCACTCTGCCCTCATGCTGTGTCAGTGGCGCTGCCCTGCCCTGCCGACGAGC 240
QY 1774 ATGATGACTTTGCTGATGACATCTCCCTGCTGAAGTGAAGAGGCCATGGCCAGAGAATA 1833
DB 241 ATGATGACTTTGCTGATGACATCTCCCTGCTGAAGTGAAGAGGCCATGGCCAGAGAATA 300
QY 1834 GAGATTCCTGTCGACACACCTCCGTTGCTACCTTTGGTCAAGTAGGAGACACAGATG 1893
DB 301 GAGATTCCTGTCGACACACCTCCGTTGCTACCTTTGGTCAAGTAGGAGACACAGATG 360
QY 1894 GCACCTGTGGCCAGACACCTTCAGGACCTCCGCCACCCAAATGCTCTGCTTGTATG 1953
DB 361 GCACCTGTGGCCAGACACCTTCAGGACCTCCGCCACCCAAATGCTCTGCTTGTATG 420
QY 1954 GAGAAGGAAAGGCTGGCAAGTGGGTTCACGGGACTGTACCTGTAGGAAACAGAAAAGA 2013
DB 421 GAGAAGGAAAGGCTGGCAAGTGGGTTCACCTGTACCTGTAGGAAACAGAAAAGA 479
QY 2014 GAAGAAGAAGCACTCTGCTGGCGGAATACTCTTGGTCACTCAATTTAAGTCGGGAA 2073
DB 480 GAAGAAGAAGCACTCTGCTGGCGGAATACTCTTGGTCACTCAATTTAAGTCGGGAA 539
QY 2074 ATTCTGCTGCTGAACCTTCAGCCCTGAACCTTTGTCCACCATTCCTTTAAATTCCTCAA 2133
DB 540 ATTCTGCTGCTGAACCTTCAGCCCTGAACCTTTGTCCACCATTCCTTTAAATTCCTCAA 599
QY 2134 CCC-AAAGTATTCCTTTCTTTAGTTTCAGAAGTACTGGCATCACACGAGGTAGCTT 2192
DB 600 CCCAAAGTATTCCTTTCTTTAGTTTCAGAAGTACTGGCATCACACGAGGTAGCTT 659
QY 2193 GGGGTGTGCTGCTGGTACCTGG 2217
DB 660 GGGGTGTGCTGCTGGTACCTGG 684

RESULT 14

BM048418 730 bp mRNA linear EST 07-NOV-2001
LOCUS
DEFINITION 603625683F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452052 5',

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

mRNA sequence.

BM048418
BM048418.1 GI:16777685
EST.
human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 730)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCMI943 row: f column: 21

High quality sequence stop: 696.

Location/Qualifiers

1. 730

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5452052"

/clone_lib="NIH_MGC_40"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

BASE COUNT 163 a 210 c 206 g 151 t

ORIGIN

Query Match 27.6%; Score 647.4; DB 13; Length 730;
Best Local Similarity 97.0%; Pred. No. 5.2e-148;
Matches 713; Conservative 0; Mismatches 16; Indels 6; Gaps 5;
QY 539 GATGACCGTGGGAGCCGCCCGGACAGCGCTCAACATCTGTTGATACAGGCACAGTAA 598
DB 2 GATGACCGTGGGAG-CCCCCGCAGACGCTCAACATCTGTTGATACAGGCACAGTAA 60
QY 599 CTTTGGCAGTGGGTGCTGCCGCCACCCCTTCTTCATCGCTACTACAGAGGAGCTGTC 658
DB 61 CTTTGGCAGTGGGTGCTGCCGCCACCCCTTCTTCATCGCTACTACAGAGGAGCTGTC 120
QY 659 CAGCACATACCGGGACCTCCGGGAAGGTGTATGTGTCCTTACACCCAGGGCAAGTGGGA 718
DB 121 CAGCACATACCGGGA-CTCCGGGAAGGTGTATGTGTCCTTACACCCA-GGCAAGTGGGA 178
QY 719 AGGGAGCTGGGACCGACCTGTTAAGATCCGCCATGGCCCCCACTGTCGCTGTC 778
DB 179 AGGGAGCTGGGACCGACCTGTTAAGATCCGCCATGG-CCCAACGCTCACTGTGCGTGC 237
QY 779 CAACATTGCTGCCATCACTGAATCAGACAAGTCTTCTATCAACGGCTCCCACTGGGAAG 838
DB 238 CAACATTGCTGCCATCACTGAATCAGACAAGTCTTCTATCAACGGCTCCCACTGGGAAG 297
QY 839 CATCTGCGGCTGGCCTATGCTGAGATTGGCAGGCGCTGACGACTCCCTGGAGCCTTTCTT 898
DB 298 CATCTGCGGCTGGCCTATGCTGAGATTGGCAGGCGCTGACGACTCCCTGGAGCCTTTCTT 357
QY 899 TGACTCTCTGTAAGAGCAGACCCACGTTCCCAACCTCTTCTCCCTGACGCTTTGTGGTGC 958
DB 358 TGACTCTCTGTAAGAGCAGACCCACGTTCCCAACCTCTTCTCCCTGACGCTTTGTGGTGC 417

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:00:56 ; Search time 85 Seconds
(without alignments)
714.850 Million cell updates/sec

Title: US-09-723-722A-43
Perfect score: 2419
Sequence: 1 ETDPEPGRGSRFVEMVD.....CLRLRQHQHDDFADISLLK 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DR seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2419	100.0	456	21 AAB07897	Active enzyme port
2	2419	100.0	488	22 AAB66572	Human memapsin 2.
3	2419	100.0	488	22 AAB61334	Memapsin 2 protein
4	2419	100.0	488	23 AAU99488	Human memapsin 2.
5	2419	100.0	501	21 AAY94767	Human beta-secreta
6	2419	100.0	501	21 AAB07896	Amino acid sequenc
7	2419	100.0	503	22 AAB66573	Human pro-memapsin
8	2419	100.0	503	22 AAB61335	T7 promoter and ve
9	2419	100.0	503	23 AAU99489	pro-memapsin 2 enc
10	2419	100.0	509	23 AAM52697	FLAG-tagged human

11	2414	99.8	501	21 AAY88425	Human aspartyl pro
12	2414	99.8	501	22 AAE10629	Human aspartyl pro
13	2414	99.8	501	22 AAE06859	Human aspartyl pro
14	2414	99.8	501	22 AAU06603	Human Aspartyl pro
15	2414	99.8	501	22 AAU07202	Human aspartyl pro
16	2414	99.8	501	22 AAE02581	Human aspartyl pro
17	2414	99.8	501	23 ABB78590	Human Asp-2(a) pro
18	2414	99.8	501	23 ABB06409	Human aspartyl pro
19	2413	99.8	501	19 AAW59807	Amino acid sequenc
20	2392	98.9	501	21 AAY94769	Rat beta-secretase
21	2390	98.8	501	21 AAY94768	Murine beta-secret
22	2390	98.8	501	21 AAY88427	Murine aspartyl pr
23	2390	98.8	501	22 AAE10631	Murine aspartyl pr
24	2390	98.8	501	22 AAE06861	Murine aspartyl pr
25	2390	98.8	501	22 AAU06605	Mouse Aspartyl pro
26	2390	98.8	501	22 AAU07204	Mouse aspartyl pro
27	2390	98.8	501	22 AAB84948	Mouse aspartic sec
28	2390	98.8	501	22 AAE02583	Murine aspartyl pr
29	2390	98.8	501	23 ABB78592	Mouse Asp-2(a) pro
30	2351.5	97.2	969	22 ABO09611	Novel human diagno
31	2320	95.9	790	19 AAW59808	Partial amino acid
32	2264.5	93.6	476	21 AAY88426	Human aspartyl pro
33	2264.5	93.6	476	22 AAE10630	Human aspartyl pro
34	2264.5	93.6	476	22 AAE06860	Human aspartyl pro
35	2264.5	93.6	476	22 AAU06604	Human Aspartyl pro
36	2264.5	93.6	476	22 AAU07203	Human aspartyl pro
37	2264.5	93.6	476	22 AAE02582	Human aspartyl pro
38	2264.5	93.6	476	23 ABB78591	Human Asp-2(b) pro
39	2264.5	93.6	476	23 ABB06410	Human aspartyl pro
40	2264.5	93.6	476	23 ABB06120	Human NS protein s
41	2238.5	92.5	476	22 AAE06909	Murine aspartyl pr
42	2238.5	92.5	476	22 AAE02619	Murine aspartyl pr
43	2212	91.4	427	22 AAY93866	Human polypeptide,
44	2160	89.3	453	21 AAY88438	Modified human asp
45	2160	89.3	453	22 AAE10642	Human-Asp 2(a) pro

ALIGNMENTS

RESULT 1
AAB07897
ID AAB07897 standard; Protein: 456 AA.
XX AAB07897;
XX AAB07897;
DT 14-NOV-2000 (first entry)
XX Active enzyme portion of human beta-secretase enzyme.
DE Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor; ss.
XX Homo sapiens.
OS WO200047618-A2.
XX
PN 17-AUG-2000.
PD
XX 10-FEB-2000; 2000WO-US03819.
PF
XX 10-FEB-1999; 99US-0119571.
PR 15-JUN-1999; 99US-0139172.
XX
XX (ELAN-) ELAN PHARM INC.
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX WPI; 2000-533011/48.
DR
XX Purified beta-secretase protein used in assays to discover inhibitors
PT


```
RESULT 3
AAB61334
ID AAB61334 standard; protein; 488 AA.
XX
AC AAB61334;
XX
DT 02-APR-2001 (first entry)
XX
DE Memapsin 2 protein.
XX
XX Memapsin 2; catalyst; Alzheimer's.
XX
OS Homo sapiens.
XX
PN W0200100663-A2.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-US17661.
XX
PR 28-JUN-1999; 99US-0141363.
XX
PR 30-NOV-1999; 99US-0168060.
XX
PR 23-JAN-2000; 2000US-0177836.
XX
PR 27-JAN-2000; 2000US-0178368.
XX
PR 08-JUN-2000; 2000US-0210292.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Tang JJN, Lin X, Koelsch G;
XX
XX WPI; 2001-102885/11.
XX
XX Purified recombinant catalytically active memapsin 2, used to screen
XX inhibitors of it, which are used to treat and prevent Alzheimer's
XX disease -
XX
XX Claim 2; Page 73-75; 86pp; English.
XX
XX The present invention relates to a purified recombinant
XX catalytically active memapsin 2. The invention may be used for
XX isolating inhibitors which are used to treat or prevent
XX Alzheimer's disease. The invention may also be used to screen
XX for individuals more genetically prone to develop Alzheimer's
XX disease.
XX
XX Sequence 488 AA;
XX
Query Match 100.0%; Score 2419; DB 22; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.2e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPPGRGGSFVEMVDNLKRGKSGQYVEMTVGSPQTLNLIVDTGSSNFVGAAP 60
DB 33 ETDEEPEEPPGRGGSFVEMVDNLKRGKSGQYVEMTVGSPQTLNLIVDTGSSNFVGAAP 92
QY 61 HPFLHRYQRLSTYRDLKRGVVPVTOGKWEGLGTLVSIPIHGNVTVRANIAITE 120
DB 93 HPFLHRYQRLSTYRDLKRGVVPVTOGKWEGLGTLVSIPIHGNVTVRANIAITE 152
QY 121 SDRKFFINGSNWEGILGLAYAEIARPDSDLRPFDSLVKQTHVPNLFSLQCGAGFPLNQS 180
DB 153 SDRKFFINGSNWEGILGLAYAEIARPDSDLRPFDSLVKQTHVPNLFSLQCGAGFPLNQS 212
QY 181 EVLASVCGSMIIGIDHSLTSGLSWYTPIRREWYEVIIIVRVEINGODLKWDCKEYNDK 240
DB 213 EVLASVCGSMIIGIDHSLTSGLSWYTPIRREWYEVIIIVRVEINGODLKWDCKEYNDK 272
QY 241 SIYDSGTTNLRPKKVEAAVKSIAASTEKPPDGFMLGEOQLVCMQAGTTPWNIFPVIS 300
DB 273 SIYDSGTTNLRPKKVEAAVKSIAASTEKPPDGFMLGEOQLVCMQAGTTPWNIFPVIS 332
QY 301 LYLMEVNTQSFRITILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV 360
DB 333 LYLMEVNTQSFRITILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV 392
QY 361 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPOTDSTLMTIAYVMAAI 420
DB 393 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPOTDSTLMTIAYVMAAI 452
QY 421 CALFMLPLCLMWCVQWRCLRLRQHQHDDFADDISLLK 456
DB 453 CALFMLPLCLMWCVQWRCLRLRQHQHDDFADDISLLK 488
RESULT 4
AAU99488
ID AAU99488 standard; Protein; 488 AA.
XX
AC AAU99488;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human memapsin 2.
XX
XX Human; memapsin 2; beta secretase; aspartic protease; APP;
XX beta-amyloid precursor protein; amyloid plaque; Alzheimer's disease;
XX neuroprotective; nootropic; enzyme.
XX
XX Homo sapiens.
XX
XX US2002049303-A1.
XX
XX 25-APR-2002.
XX
XX 28-FEB-2001; 2001US-0796264.
XX
XX 28-JUN-1999; 99US-141363P.
XX
XX 30-NOV-1999; 99US-168060P.
XX
XX 25-JAN-2000; 2000US-177836P.
XX
XX 27-JAN-2000; 2000US-178368P.
XX
XX 27-JUN-2000; 2000US-0604608.
XX
XX (TANG/) TANG J J N.
XX (LINX/) LIN X.
XX (KOEL/) KOELSCH G.
XX (HONG/) HONG L.
XX
XX Tang JJN, Lin X, Koelsch G, Hong L;
XX
XX WPI; 2002-507280/54.
XX
XX N-PSDB; ABK88641.
XX
XX New recombinant catalytically active memapsin 2, useful to screen for
XX inhibitors of memapsin 2 which can be used to prevent and treat
XX Alzheimer's disease
XX
XX Claim 2; Page 22-23; 44pp; English.
XX
XX The present invention relates to methods for the production of
XX purified, recombinant catalytically active, memapsin 2 (beta
XX secretase). Memapsin 2, a member of the aspartic protease family,
XX cleaves beta-amyloid precursor protein (APP) found in amyloid plaques.
XX The recombinant memapsin 2 is useful for identifying inhibitors of
XX memapsin 2 in the design of drugs for the treatment and/or prevention
XX of Alzheimer's disease. The recombinant memapsin 2 can be used to
XX immunise against Alzheimer's disease. The present sequence represents
XX human memapsin 2.
XX
XX Sequence 488 AA;
Query Match 100.0%; Score 2419; DB 23; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.2e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPPGRGGSFVEMVDNLKRGKSGQYVEMTVGSPQTLNLIVDTGSSNFVGAAP 60
```

Db 33 ETDEPEEPGRGSGFVEMVDNLGRSGQGYVEMTVCSPPTLNILVDTGSSNFAVGAAP 92
QY 61 HPFLHRYQROLSSYRDLRGVYVYPTQKGWEGELGTLVSPHGPNTVVRANIAATE 120
Db 93 HPFLHRYQROLSSYRDLRGVYVYPTQKGWEGELGTLVSPHGPNTVVRANIAATE 152
QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDSLEPPFDSLVKQTHVFNLFSLQLCGAGFPLNQS 180
Db 153 SDRFFINGSNWEGILGLAYAEIARPDSDSLEPPFDSLVKQTHVFNLFSLQLCGAGFPLNQS 212
QY 181 EVLASVCGSMIIGGIDHSLYTGSLWYTPTRREWYEVIIIVRVEINGODLKMCKEYNYDK 240
Db 213 EVLASVCGSMIIGGIDHSLYTGSLWYTPTRREWYEVIIIVRVEINGODLKMCKEYNYDK 272
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKFPDGFGLGEQLVCMWAGTTPWNIFFVIS 300
Db 273 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKFPDGFGLGEQLVCMWAGTTPWNIFFVIS 332
QY 301 LYLMEVNTNQSFRITILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 333 LYLMEVNTNQSFRITILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 392
QY 361 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
Db 393 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 452
QY 421 CALFMLPLCLMVCWRCRLRQHQHDDFADDISLLK 456
Db 453 CALFMLPLCLMVCWRCRLRQHQHDDFADDISLLK 488

RESULT 5
ID AAY94767 standard; Protein: 501 AA.
AC AAY94767;
DT 12-FEB-2001 (first entry)
XX Human beta-secretase amino acid sequence.
DE Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; human;
KW Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..45
FT Protein /label= putative signal peptide
FT 46..501
FT /label= Beta-secretase
XX WO200058479-A1.
XX 05-OCT-2000.
XX 23-MAR-2000; 2000WO-US07755.
XX 26-MAR-1999; 99US-0277229.
XX (AMGE-) AMGEN INC.
XX Citron M, Vassar RJ, Bennett BD;
XX WPI; 2000-594643/56.
XX N-PSDB; AAA28278.
XX Isolated beta-secretase' nucleic acids and encoded polypeptides, useful
PT for diagnosis and gene therapy of Alzheimer's disease
XX
XX Claim 1; Fig 4; 145pp; English.

CC This invention relates to 3 nucleotide sequences encoding beta-secretase
CC proteins. Beta-secretase is an enzyme involved in the production of one
CC of the components of amyloid plaques involved in Alzheimer's disease. The
CC invention includes an expression vector comprising the nucleotide
CC sequence, a host cell comprising the expression vector, and a process for
CC producing the protein through culturing the transformed cells. Also
CC included in the invention are a polypeptide derivative of the
CC beta-secretase protein, a fusion protein comprising beta-secretase fused
CC to a heterologous amino acid sequence, and a method for modulating the
CC levels of beta-secretase polypeptide in a mammal comprising administering
CC the polynucleotide sequence. Beta-secretase exhibits neuroprotective and
CC neurotropic activity. The beta-secretase nucleotide sequence may be used to
CC map locations of the beta-secretase gene and related genes on chromosomes
CC and as hybridization probes in diagnostic assays to test for the presence
CC of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's
CC syndrome, and amyloid angiopathy. The nucleotide sequence may also be
CC used as anti-sense inhibitors of beta-secretase expression, in gene
CC therapy of Alzheimer's disease, and for the identification of compounds
CC that modulate beta-secretase activity. Antibodies to the beta-secretase
CC protein may be used for in vitro and in vivo diagnostic purposes to
CC detect the presence of beta-secretase polypeptide in a body fluid or cell
CC sample. The present sequence represents the human beta-secretase protein.
XX
SQ Sequence 501 AA:
Query Match 100.0%; Score 2419; DB 21; Length 501;
Best Local Similarity 100.0%; Pred. No. 5,5e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEPEEPGRGSGFVEMVDNLGRSGQGYVEMTVCSPPTLNILVDTGSSNFAVGAAP 60
Db 46 ETDEPEEPGRGSGFVEMVDNLGRSGQGYVEMTVCSPPTLNILVDTGSSNFAVGAAP 105
QY 61 HPFLHRYQROLSSYRDLRGVYVYPTQKGWEGELGTLVSPHGPNTVVRANIAATE 120
Db 106 HPFLHRYQROLSSYRDLRGVYVYPTQKGWEGELGTLVSPHGPNTVVRANIAATE 165
QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDSLEPPFDSLVKQTHVFNLFSLQLCGAGFPLNQS 180
Db 166 SDRFFINGSNWEGILGLAYAEIARPDSDSLEPPFDSLVKQTHVFNLFSLQLCGAGFPLNQS 225
QY 181 EVLASVCGSMIIGGIDHSLYTGSLWYTPTRREWYEVIIIVRVEINGODLKMCKEYNYDK 240
Db 226 EVLASVCGSMIIGGIDHSLYTGSLWYTPTRREWYEVIIIVRVEINGODLKMCKEYNYDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKFPDGFGLGEQLVCMWAGTTPWNIFFVIS 300
Db 286 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKFPDGFGLGEQLVCMWAGTTPWNIFFVIS 345
QY 301 LYLMEVNTNQSFRITILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 346 LYLMEVNTNQSFRITILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 405
QY 361 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
Db 406 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMVCWRCRLRQHQHDDFADDISLLK 456
Db 466 CALFMLPLCLMVCWRCRLRQHQHDDFADDISLLK 501
RESULT 6
ID AAB07896
XX AAB07896 standard; Protein: 501 AA.
XX AC AAB07896;
XX 14-NOV-2000 (first entry)
XX Amino acid sequence of a human beta-secretase enzyme.
DE Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW

KW anyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.

OS Homo sapiens.

PN WO200047618-A2.

XX 17-AUG-2000.

XX 10-FEB-2000; 2000WO-US03819.

XX 10-FEB-1999; 99US-0119571.

PR 15-JUN-1999; 99US-0139172.

XX (ELAN-) ELAN PHARM INC.

XX Anderson JP, Basi G, Doane MT, Erigon N, John V, Power M;

PI Sinh S, Tatsuno G, Tung J, Wang J, McConlogue L;

XX WPI; 2000-533011/48.

DR N-PSDB; AAA59550, AAA59551.

XX Purified beta-secretase protein used in assays to discover inhibitors

PT which can be used for the treatment of amyloidogenic diseases e.g.

XX Alzheimer's disease -

XX Claim 17; Fig 2A; 121pp; English.

XX The specification describes a beta-secretase enzyme. The enzyme cleaves

CC beta-amyloid precursor protein to produce beta-amyloid peptide. This

CC enzyme is therefore implicated in the production of amyloid plaque

CC components which accumulate in the brains of individuals afflicted with

CC Alzheimer's disease. Inhibitors of beta-secretase are administered to

CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's

CC disease-like pathology to test if they maintain or improve cognitive

CC ability or reduce the plaque burden. The compounds are used for the

CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The

CC present sequence represents a human beta-secretase enzyme.

XX Sequence 501 AA;

Query Match 100.0%; Score 2419; DB 21; Length 501;

Best Local Similarity 100.0%; Pred. No. 5.5e-242;

Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGSGFVEMVDNLKSGQYVEMTVGSPPTLNLVDTGSSNFAVGAAP 60

DB 46 ETDEEPEEPRGSGFVEMVDNLKSGQYVEMTVGSPPTLNLVDTGSSNFAVGAAP 105

QY 61 HPFLHRYQRLSSTYRDLRKGVPVPTQGWEGELGTDLSVIPHGPNTVVRANIAATE 120

DB 106 HPFLHRYQRLSSTYRDLRKGVPVPTQGWEGELGTDLSVIPHGPNTVVRANIAATE 165

QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDLLEPFFDSLKVQTHVPLNLSLQLCGAGFPLNQ 180

DB 166 SDRFFINGSNWEGILGLAYAEIARPDSDLLEPFFDSLKVQTHVPLNLSLQLCGAGFPLNQ 225

QY 181 EVLASVCGSMITGIDHSLTSGSLWYTPIRREWYVEVILVRVEINGODLKMCKEYNVNDK 240

DB 226 EVLASVCGSMITGIDHSLTSGSLWYTPIRREWYVEVILVRVEINGODLKMCKEYNVNDK 285

QY 241 SIYDSGTTNLRPKKVFEEAAVKSIKAASTTEKFPDGFGLGQVLCVQAGTTPWNIFPVIS 300

DB 286 SIYDSGTTNLRPKKVFEEAAVKSIKAASTTEKFPDGFGLGQVLCVQAGTTPWNIFPVIS 345

QY 301 LYLMEVNTNSFRITILPQOYLRPEVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYV 360

DB 346 LYLMEVNTNSFRITILPQOYLRPEVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYV 405

QY 361 FDRARKRIGFAVSACHVHDEFTAAVGPFTVLDMECGYNIPQDDESTLMTIAYVMAAI 420

DB 406 FDRARKRIGFAVSACHVHDEFTAAVGPFTVLDMECGYNIPQDDESTLMTIAYVMAAI 465

QY 421 CALFMLPLCLMVCQWRCRLCRLRQHQHDDFADDISLLK 456
DB 466 CALFMLPLCLMVCQWRCRLCRLRQHQHDDFADDISLLK 501

RESULT 7
AAB66573

ID AAB66573 standard; Protein; 503 AA.

XX AAB66573;

XX 12-APR-2001 (first entry)

DE Human pro-memapsin 2.

KW Human; memapsin 2; nootropic; neuroprotective; amyloid precursor protein;

KW APP; memapsin 2 inhibitor; Alzheimer's disease; ss.

OS Homo sapiens.

PN WO200100665-A2.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-US17742.

XX 28-JUN-1999; 99US-01411363.

XX 30-NOV-1999; 99US-0168060.

XX 25-JAN-2000; 2000US-0177836.

XX 27-JAN-2000; 2000US-0178368.

XX 08-JUN-2000; 2000US-0210292.

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX (UNII) UNIV ILLINOIS FOUND.

XX Tang JJN, Hong L, Ghosh AK;

XX WPI; 2001-137933/14.

XX Novel memapsin 2 inhibitors which bind to active site of memapsin 2

PT having 2 catalytic aspartic residues and substrate binding cleft, used

PT to treat Alzheimer's disease by blocking amyloid precursor protein

PT cleavage -

XX Example 4; Fig 1; 86pp; English.

PS The present sequence is given in a specification relating to an inhibitor

CC of catalytically active memapsin 2. The inhibitor binds to the memapsin 2

CC active site, which is defined by the presence of two catalytic aspartic

CC residues and a substrate binding cleft. The inhibitor is useful for

CC the treatment and diagnosis of Alzheimer's disease. It is useful in

CC screens for individuals with a genetic predisposition to Alzheimer's

CC disease. The inhibitor is useful as a reagent for specifically binding to

CC memapsin 2 or memapsin 2 analogues and for aiding in memapsin 2

CC isolation, purification and characterisation.

XX Sequence 503 AA;

Query Match 100.0%; Score 2419; DB 22; Length 503;

Best Local Similarity 100.0%; Pred. No. 5.5e-242;

Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGSGFVEMVDNLKSGQYVEMTVGSPPTLNLVDTGSSNFAVGAAP 60

DB 48 ETDEEPEEPRGSGFVEMVDNLKSGQYVEMTVGSPPTLNLVDTGSSNFAVGAAP 107

QY 61 HPFLHRYQRLSSTYRDLRKGVPVPTQGWEGELGTDLSVIPHGPNTVVRANIAATE 120

DB 108 HPFLHRYQRLSSTYRDLRKGVPVPTQGWEGELGTDLSVIPHGPNTVVRANIAATE 167

QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDLLEPFFDSLKVQTHVPLNLSLQLCGAGFPLNQ 180

DB 168 SDRFFINGSNWEGILGLAYAEIARPDSDLLEPFFDSLKVQTHVPLNLSLQLCGAGFPLNQ 227

[illegible]

RESULT 8
AAB61335
ID AAB61335 standard; protein; 503 AA.

PT Purified recombinant catalytically active memapsin 2, used to screen
PT inhibitors of it, which are used to treat and prevent Alzheimer's
PT disease -

AA The present invention relates to a purified recombinant
CC catalytically active memapsin 2. The invention may be used for
CC isolating inhibitors which are used to treat or prevent
CC Alzheimer's disease. The invention may also be used to screen
CC for individuals more genetically prone to develop Alzheimer's
CC disease.

Query Match 100.0%; Score 2419; DB 22; Length 503;
Best Local Similarity 100.0%; pred. NO. 5.5e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0;

RESULT 9
AAU99489
ID AAU99489
XX
AC AAU99489
XX
DT 07-07-2011

Qy	1	ETDEPEPEGRGSGFVEMVDNLURKSGQGYVEMT	VGSPPTLNI	LVDTGSSNF	AVGAAP	60
Db	48	ETDEPEPEGRGSGFVEMVDNLURKSGQGYVEMT	VGSPPTLNI	LVDTGSSNF	AVGAAP	107
Qy	61	HPFLHRYTORQLSSYRDLRKGVYVPTQCKWEGELGTDLVS	IPHGPNVTVRANIAA	ITE	120	
Db	108	HPFLHRYTORQLSSYRDLRKGVYVPTQCKWEGELGTDLVS	IPHGPNVTVRANIAA	ITE	167	
Qy	121	SDKEFFINGSNWEGILGLAYAEIARPDSDLEPPFDSL	VKQTHVPNL	FSLQLCGAG	PPLNQS	180
Db	168	SDKEFFINGSNWEGILGLAYAEIARPDSDLEPPFDSL	VKQTHVPNL	FSLQLCGAG	PPLNQS	227
Qy	181	EVLASVGSGMIICGIDHSLYTGS	LWTPPIRREWY	EVI	I	240
Db	228	EVLASVGSGMIICGIDHSLYTGS	LWTPPIRREWY	EVI	I	287
Qy	241	SIVDSGTTNLRLPKPKVFEEAAVKS	IKAASST	ETKFPDGF	EWLGQV	300
Db	288	SIVDSGTTNLRLPKPKVFEEAAVKS	IKAASST	ETKFPDGF	EWLGQV	347
Qy	301	LYLMGEVNTQSPRITILPOQYLR	PVEDVAT	SODDCKY	KAISQSS	360
Db	348	LYLMGEVNTQSPRITILPOQYLR	PVEDVAT	SODDCKY	KAISQSS	407
Qy	361	FDRAKRIGFAVSACHVHDEE	FRTAAVEG	FPVTLDM	EDCGYNI	420
Db	408	FDRAKRIGFAVSACHVHDEE	FRTAAVEG	FPVTLDM	EDCGYNI	467
Qy	421	CALFMLPLCLMVCWRCRLC	LROQHD	FDADD	ISLLK	456
Db	468	CALFMLPLCLMVCWRCRLC	LROQHD	FDADD	ISLLK	503

RESULT 9
AAU99489
ID AAU99489 standard; Protein: 503 AA.
XX
XX AAU99489;
XX AC
XX AC
DT 07-OCT-2002 (first entry)

XX	
DR	
XX	WPI; 2002-507280/54.
PT	New recombinant catalytically active memapsin 2, useful to screen for

Db	226	EVLASVGGSMIIIGIDHSlyTGSlyWYTPiRREWYVEVlIiVRVEiNGQDLKMDCKEYNYDK	28
Qy	241	SIVDSGTTNLRLPKKVFEEAAVKSIKAASSTeKfPDGfWLGbOLVCWQAGTTPWNlFPVlS	300
Db	286	SIVDSGTTNLRLPKKVFEEAAVKSIKAASSTeKfPDGfWLGbOLVCWQAGTTPWNlFPVlS	345
Qy	301	LYLMGEVtNQSFRIITiLPQOYLRPVEDVATsQDDCYKfAIsOSSTGTVMGAVIMEGFYVW	360
Db	346	LYLMGEVtNQSFRIITiLPQOYLRPVEDVATsQDDCYKfAIsOSSTGTVMGAVIMEGFYVW	405
Qy	361	FDRARKRIGfAVsACHVHDEfRTAAVEGPFVTLDMEDCGYNIPOtDESTLMTiAYVMAAI	420
Db	406	FDRARKRIGfAVsACHVHDEfRTAAVEGPFVTLDMEDCGYNIPOtDESTLMTiAYVMAAI	465
Qy	421	CALFMPLPLCLMVCWRCLRCLRQOHDDfADDISLLK	456
Db	466	CALFMPLPLCLMVCWRCLRCLRQOHDDfADDISLLK	501
RESULT 12			
ID	AAE10629		
XX	AAE10629	standard; Protein; 501 AA.	
AC	AAE10629;		
XX			
DT	10-DEC-2001	(first entry)	
XX			
DE	Human aspartyl	protease 2(a) [hu-Asp2(a)] protein.	
XX			
KW	Human; aspartyl	protease 2(a); Asp2(a); amyloid precursor protein; API	
KW	Alzheimer's	disease; AD; dementia; neurofibrillary tangle; gliosis;	
KW	amyloid plaque; neuronal	loss; proteolytic; nontropic; neuroprotective	
XX	chromosome	11q23.3-24.1.	
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..21	
FT	Peptide	/label= Signal_peptide	
FT	Peptide	22..45	
FT	Peptide	/label= Asp_2a_prepropeptide	
FT	Peptide	46..57	
FT	Protein	/label= Asp_2a_propeptide	
FT	Protein	58..501	
FT	Region	/label= Mature_human_Asp_2a_protein	
FT	Region	420..454	
FT	Domain	/label= Alpha-helical_spacer_region	
FT	Domain	455..477	
FT	Domain	/label= Transmembrane_domain	
FT	Domain	478..501	
XX		/label= Cytoplasmic_domain	
PN	GB2357767-A.		
XX			
PD	04-JUL-2001.		
XX			
PF	22-SEP-2000;	2000GB-0023315.	
XX			
PR	23-SEP-1999;	99US-0155493.	
PR	23-SEP-1999;	99US-0404133.	
PR	23-SEP-1999;	99WO-US20881.	
PR	13-OCT-1999;	99US-0416901.	
PR	06-DEC-1999;	99US-0169232.	
XX			
PA	(PHAA) PHARMACIA & UPJOHN CO.		
XX			
PI	Bienkowski MJ, Gurney M;		
XX			
DR	WPI: 2001-444208/48.		
DR	N-PSDB; AAD17865.		
XX			
PT	Polypeptide comprising fragments of human aspartyl	protease with	
PT	amyloid precursor protein processing activity and alpha-secretase		

PT activity, for identifying modulators useful in treating Alzheimer's
PT disease -
PS Example 2; Fig 2; 187pp; English.
XX The patent discloses human aspartyl protease 1 (hu-Asp1) or modified
CC Asp1 proteins which lack transmembrane domain or amino terminal
CC domain or cytoplasmic domain and retains alpha-secretase activity
CC and amyloid protein precursor (APP) processing activity. The proteins
CC of the invention are useful for assaying hu-Asp1 alpha-secretase
CC activity, which in turn is useful for identifying modulators of
CC hu-Asp1 alpha-secretase activity, where modulators that increase
CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
CC disease (AD) which causes progressive dementia with consequent
CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying
CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
CC the substrate under acidic conditions and determining the level of
CC hu-Asp1 proteolytic activity. The present sequence is long form of
CC human Asp2 protein, designated as Asp2(a). Asp2 gene is localised
CC on chromosome 11q23.3-24.1.
XX
SQ Sequence 501 AA;
Query Match 99.8%; Score 2414; DB 22; Length 501;
Best Local Similarity 99.8%; Pred. No. 1.8e-241;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ETDEEPEEPPGRGSGFVEMVDNLKSGQGYVYVEMTSGSPPTLNILVDTGSSNFAVGAAP 60
DB 46 ETDEEPEEPPGRGSGFVEMVDNLKSGQGYVYVEMTSGSPPTLNILVDTGSSNFAVGAAP 105
QY 61 HPFLHRYRQQLSSTYRDLKRGVVPYVPTQKWECELGTDLSVPHGPNVTVRANIAITE 120
DB 106 HPFLHRYRQQLSSTYRDLKRGVVPYVPTQKWECELGTDLSVPHGPNVTVRANIAITE 165
QY 121 SDRFFINGSNWEGILGLAYAEIARPDLSLEPFDLSLVKQTHVPLNLSLQCGAGFPLNQS 180
DB 166 SDRFFINGSNWEGILGLAYAEIARPDLSLEPFDLSLVKQTHVPLNLSLQCGAGFPLNQS 225
QY 181 EVLASVGGSMIIGIDHSLTGLSWYTPIRREWYEVYIIVRVEINGQDLKMDCKEYNDK 240
DB 226 EVLASVGGSMIIGIDHSLTGLSWYTPIRREWYEVYIIVRVEINGQDLKMDCKEYNDK 285
QY 241 SIYDSGTTNLRPKKVEAAVKSIIKAASSTKFPDGFMLGEQLVCWAGTTPWNIFPVIS 300
DB 286 SIYDSGTTNLRPKKVEAAVKSIIKAASSTKFPDGFMLGEQLVCWAGTTPWNIFPVIS 345
QY 301 LYLMEVNTQSFRTILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 360
DB 346 LYLMEVNTQSFRTILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 405
QY 361 FDRARRIGFAVSACHVHDEFTAAVGGPVFTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DB 406 FDRARRIGFAVSACHVHDEFTAAVGGPVFTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMWQWRCRLCRLRQOHDFFADDISLLK 456
DB 466 CALFMLPLCLMWQWRCRLCRLRQOHDFFADDISLLK 501
RESULT 13
AAE06859
ID AAE06859 standard; Protein; 501 AA.
XX
AC AAE06859;
XX
DT 23-OCT-2001 (first entry)
XX
DE Human aspartyl protease 2a (Hu-Asp2a) protein.
XX Human; aspartyl protease 2a; Asp 2a; beta-amyloid precursor protein; APP;
KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
KW

KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nootropic;
KW neuroprotective; antisense therapy; gene therapy;
KW chromosome 11q23.3-24.1.
XX
OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= Signal_peptide
FT Protein 22..501
FT /note= "Mature human aspartyl protease 2a (Hu-Asp2a)"
FT Region 420..454
FT /note= "Alpha helical spacer region"
FT Domain 455..477
FT /label= Transmembrane_domain
FT Domain 478..501
FT /label= Cytoplasmic_domain
XX W0200150829-A2.
PN 19-JUL-2001.
XX
PD 09-MAY-2001; 2001WO-IB00799.
PF 09-MAY-2001; 2001WO-IB00799.
XX
PR (BIEN/) BIENKOWSKI M J.
PA (GURW/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX WPI; 2001-483072/52.
DR N-PSDB; AAD13021.
XX
XX Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity
XX
PS Claim 49; Fig 2; 185pp; English.
XX
CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
CC precursor protein (APP) isoforms and their corresponding DNA molecules.
CC Human aspartyl proteases can act as beta-secretase proteases useful for
CC treating Alzheimer's disease. APP isoforms are useful for identifying
CC modulators of amyloid-beta peptide production, for use in designing
CC therapeutics for the treatment and prevention of Alzheimer's disease,
CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
CC and neuronal loss. APP isoforms are also used in methods for identifying
CC inhibitors and modulators of human Asp2 activity. The invention relates
CC to a method for identifying agents that modulate the activity of human
CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
CC as a means to screen in cellular assays for the inhibitors of beta- and
CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
CC polymerase chain reactions (PCR). The probes are useful for detecting
CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
CC blots. The present sequence is human aspartyl protease 2 (Hu-Asp2), a
CC 'long' form designated as (Hu-Asp2a). Hu-Asp 2 gene is localised on
CC chromosome 11q23.3-24.1.
XX
SQ Sequence 501 AA;
Query Match 99.8%; Score 2414; DB 22; Length 501;
Best Local Similarity 99.8%; Pred. No. 1.8e-241;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ETDEEPEEPPGRGSGFVEMVDNLKSGQGYVYVEMTSGSPPTLNILVDTGSSNFAVGAAP 60
DB 46 ETDEEPEEPPGRGSGFVEMVDNLKSGQGYVYVEMTSGSPPTLNILVDTGSSNFAVGAAP 105

QY 61 HPFLHRYQRLSTYRDLRGVYVPYTGQKWEGLGTHDLSVPHGPNVTVRANIAAITE 120
Db 106 HPFLHRYQRLSTYRDLRGVYVPYTGQKWEGLGTHDLSVPHGPNVTVRANIAAITE 165
QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFEDSLVKQTHVPNLFSLHLCGAGFPLNQS 180
Db 166 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFEDSLVKQTHVPNLFSLHLCGAGFPLNQS 225
QY 181 EVLASVGGSMIIGDHSLSYTGSLWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 240
Db 226 EVLASVGGSMIIGDHSLSYTGSLWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIAASSTKFPDGFMLGEOVCWQAGTTPWNIFPVIS 300
Db 286 SIVDSGTTNLRPKKVFEEAAVKSIAASSTKFPDGFMLGEOVCWQAGTTPWNIFPVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYV 360
Db 346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYV 405
QY 361 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 420
Db 406 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMVCMQWRCLRCLROQHDDFADDDISLLK 456
Db 466 CALFMLPLCLMVCMQWRCLRCLROQHDDFADDDISLLK 501

RESULT 14

AAU06603
ID AAU06603 standard; Protein: 501 AA.
AC AAU06603;
XX
XX
DT 24-OCT-2001 (first entry)
XX
DE Human Aspartyl protease 2(a), Asp2(a).
XX
KW Human; Aspartyl protease; Asp2(a); beta-secretase; neurotropic;
KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
KW amyloid-beta; Abeta.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..21
FT Peptide /label= Signal peptide
FT Peptide 22..45
FT Peptide /label= Pre_pro-peptide
FT Peptide 46..57
FT Protein /label= Pro-peptide
FT Protein 57..501
FT Region /label= Mature_Asp2(a)
FT Region 420..454
FT Domain /label= Alpha_helical_spacer_region
FT Domain 455..477
FT Domain /label= Transmembrane_domain
FT Domain 478..501
FT Domain /label= Cytoplasmic_domain

PN WO200149098-A2.

XX
XX 12-JUL-2001.

XX
XX 09-MAY-2001; 2001WO-IB00798.

XX
XX 09-MAY-2001; 2001WO-IB00798.

XX (BIEN/) BIENKOWSKI M J.
XX (GURN/) GURNEY M E.
XX (HEIN/) HEINRIKSON R L.
XX (PARO/) PARODI L A.

(YANR/) YAN R.
XX
XX Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Van R;
DR WPT: 2001-502549/55.
XX N-PSDB: AAS11517.

Novel purified polypeptide comprising fragment of mammalian aspartyl
protease 2, lacking Asp2 transmembrane domain and retaining beta
secretase activity of Asp2 useful for identifying inhibitors of Asp2
activity

Claim 49; Fig 2; 185pp; English.

The invention relates to a purified polypeptide comprising a fragment of
mammalian aspartyl protease (Asp2) protein which lacks the Asp2
transmembrane domain and the Asp2 protein, and where the polypeptide and
the fragment retain the beta-secretase activity of the mammalian Asp2
protein. The invention also details polynucleotides for the Asp
proteins and vectors expressing them, and a polypeptide (isoform of
amyloid protein precursor (APP)) comprising the amino acid sequence of an
APP or its fragment containing an APP cleavage site recognizable by a
mammalian beta-secretase, and further comprising two lysine residues at
the carboxyl terminus of the amino acid sequence of the mammalian APP or
APP fragment. Also included in the invention are methods of identifying
modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
useful for treating Alzheimer's disease. APP is useful in methods for
identifying inhibitors or modulators of human Asp2 activity and
amyloid-beta (Abeta) peptide production. APP is also useful in designing
therapeutics for the treatment or prevention of Alzheimer's disease.
APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
is associated with increased levels of Abeta processing is useful in
assays relating the Alzheimer's research. The expression vector is useful
for recombinantly expressing APP. Nucleic acids that hybridize to
Asp oligonucleotides are useful as probes or primers. The probes are
useful for detecting Hu-Asp nucleic acids in in vitro assays and in
Northern and Southern blots. The present sequence is human Asp2(a).

Sequence 501 AA:

Query Match 99.8%; Score 2414; DB 22; Length 501;
Best Local Similarity 99.8%; Pred. NO. 1.8e-241;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEPEPEGRGSGFVEMVDNLRGKSGQGYVEMTVGSPOTLNLVDTGSSNFAVGAAP 60
Db 46 ETDEPEPEGRGSGFVEMVDNLRGKSGQGYVEMTVGSPOTLNLVDTGSSNFAVGAAP 105
QY 61 HPFLHRYQRLSTYRDLRGVYVPYTGQKWEGLGTHDLSVPHGPNVTVRANIAAITE 120
Db 106 HPFLHRYQRLSTYRDLRGVYVPYTGQKWEGLGTHDLSVPHGPNVTVRANIAAITE 165
QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFEDSLVKQTHVPNLFSLHLCGAGFPLNQS 180
Db 166 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFEDSLVKQTHVPNLFSLHLCGAGFPLNQS 225
QY 181 EVLASVGGSMIIGDHSLSYTGSLWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 240
Db 226 EVLASVGGSMIIGDHSLSYTGSLWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 285
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIAASSTKFPDGFMLGEOVCWQAGTTPWNIFPVIS 300
Db 286 SIVDSGTTNLRPKKVFEEAAVKSIAASSTKFPDGFMLGEOVCWQAGTTPWNIFPVIS 345
QY 301 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYV 360
Db 346 LYLMEVNTNOSFRITILPQOYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYV 405
QY 361 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 420
Db 406 FDRARKRIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 465
QY 421 CALFMLPLCLMVCMQWRCLRCLROQHDDFADDDISLLK 456

Db 466 CALFMLPLCLMWQWRCRLRQHQHDDFADDISLLK 501

RESULT 15

AAU07202
ID AAU07202 standard; Protein: 501 AA.

XX AC AAU07202;

XX DT 24-OCT-2001 (first entry)

XX DE Human aspartyl protease 2a (Asp-2a).

XX KW Human; aspartyl protease 1; Asp-1; nootropic; neuroprotective;

XX KW aspartyl protease 2; Asp2; amyloid protein precursor; APP;

XX KW beta-secretase; Alzheimer's disease.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..21 "Signal peptide"

FT Misc_feature 22..45 "Pre-propeptide"

FT Misc_feature 46..57 "Propeptide"

FT Protein 58..501

FT Region 420..454 "Mature Aspartyl protease-2a"

FT Domain 455-477 "Alpha helical spacer region"

FT Domain 478..501 "Transmembrane domain"

FT Domain 478..501 "Cytoplasmic domain"

XX WO200149097-A2.

XX PD 12-JUL-2001.

XX PF 09-MAY-2001; 2001WO-IB00797.

XX PR 09-MAY-2001; 2001WO-IB00797.

XX PA (BIEN/) BIENKOWSKI M J.

XX PA (GURN/) GURNEY M E.

XX PA (HEIN/) HEINRIKSON R L.

XX PA (PARO/) PARODI L A.

XX PA (YANR/) YAN R.

XX PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

XX WPI: 2001-502548/55.

XX DR N-PSDB: AAS11702.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl

XX protease 2, lacking Asp2 transmembrane domain and retaining beta

XX secretase activity of Asp2 useful for identifying inhibitors of Asp2

XX activity -

XX Claim 49; Fig 2; 185pp; English.

XX The invention relates to a novel purified polypeptide comprising a

XX fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the

XX Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide

XX and the fragment retain the beta-secretase activity of the mammalian Asp2

CC of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that
CC modulate the activity of Asp2; and for reducing cellular production of
CC amyloid beta (Abeta) from APP. Agents identified by the above methods
CC are useful for treating Alzheimer's disease; and for identifying
CC modulators of amyloid-beta (Abeta) peptide production, for use in
CC designing therapeutics for the treatment or prevention of Alzheimer's
CC disease. Probes and primers derived from Asp nucleic acid sequences
CC are useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC Northern and Southern blots. The present sequence represents the
CC amino acid sequence of human Asp-2a used in the methods of the invention.
XX

SQ Sequence 501 AA;

Query Match 99.8%; Score 2414; DB 22; Length 501;
Best Local Similarity 99.8%; Pred No. 1.8e-241;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGRGSFVEMVDNLRGKSGQGYVEMTVGSPPTNLILVDTGSSNFAVGAAP 60
DB 46 ETDEEPEEPRGRGSFVEMVDNLRGKSGQGYVEMTVGSPPTNLILVDTGSSNFAVGAAP 105
QY 61 HPFLHRYRQRLSSYRDLRKGVYVPTQGWEGELGTLVSPHGPNTVTRANIAAITE 120
DB 106 HPFLHRYRQRLSSYRDLRKGVYVPTQGWEGELGTLVSPHGPNTVTRANIAAITE 165
QY 121 SDKFFINGSNWEGILGLAYAEIARPDLSLEPRFDSLKQTHVPLNLSLQLCGAGFPLNQ 180
DB 166 SDKFFINGSNWEGILGLAYAEIARPDLSLEPRFDSLKQTHVPLNLSLQLCGAGFPLNQ 225
QY 181 EVLASVGGSMIIGGIDHSYLTGSLWYTPIRREWYEVIIVRVEINGDLKMDCKEYNYDK 240
DB 226 EVLASVGGSMIIGGIDHSYLTGSLWYTPIRREWYEVIIVRVEINGDLKMDCKEYNYDK 285
QY 241 SIVDSGTTNLRPLPKVFEAAVKSIAASSTKEKPDGFWLGEQLVCWQAGTTPWNIIPV 300
DB 286 SIVDSGTTNLRPLPKVFEAAVKSIAASSTKEKPDGFWLGEQLVCWQAGTTPWNIIPV 345
QY 301 LYLMEVNTQSPRITILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 360
DB 346 LYLMEVNTQSPRITILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 405
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESLMTIAYVMAAI 420
DB 406 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESLMTIAYVMAAI 465
QY 421 CALFMLPLCLMWQWRCRLRQHQHDDFADDISLLK 456
DB 466 CALFMLPLCLMWQWRCRLRQHQHDDFADDISLLK 501

Search completed: March 4, 2003, 10:03:45

Job time : 88 secs

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	2419	100.0	488	9	US-09-795-903A-2		Sequence 2, Appli
2	2419	100.0	488	10	US-09-796-264-2		Sequence 2, Appli
3	2419	100.0	488	10	US-09-845-226-2		Sequence 2, Appli
4	2419	100.0	503	9	US-09-795-903A-3		Sequence 3, Appli
5	2419	100.0	503	10	US-09-796-264-3		Sequence 3, Appli
6	2419	100.0	503	10	US-09-845-226-3		Sequence 3, Appli
7	2414	99.8	501	10	US-09-794-927-4		Sequence 4, Appli
8	2414	99.8	501	10	US-09-795-847-4		Sequence 4, Appli
9	2414	99.8	501	10	US-09-794-743-4		Sequence 4, Appli
10	2414	99.8	501	10	US-09-794-748-4		Sequence 4, Appli
11	2414	99.8	501	10	US-09-794-925-4		Sequence 4, Appli
12	2414	99.8	501	10	US-09-681-442-4		Sequence 4, Appli
13	2413	99.8	501	9	US-09-969-671A-2		Sequence 2, Appli
14	2390	98.8	501	10	US-09-794-927-8		Sequence 8, Appli
15	2390	98.8	501	10	US-09-795-847-8		Sequence 8, Appli
16	2390	98.8	501	10	US-09-794-743-8		Sequence 8, Appli
17	2390	98.8	501	10	US-09-794-748-8		Sequence 8, Appli
18	2390	98.8	501	10	US-09-794-925-8		Sequence 8, Appli
19	2390	98.8	501	10	US-09-681-442-8		Sequence 8, Appli

OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,
OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
OTHER INFORMATION: and 427-431 are C-lobe Helices
US-09-795-903A-2

Query Match 100.0%; Score 2419; DB 9; Length 488;
Best Local Similarity 100.0%; Pred. No. 2.3e-224;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEPEGRGSEVEMVDNLGRKSGQGYVEMTSGSPQTLNLLVDTGSSNFAYGAAP 60
DQ 33 ETDEPEPEGRGSEVEMVDNLGRKSGQGYVEMTSGSPQTLNLLVDTGSSNFAYGAAP 92
QY 61 HPFLHRYQRLSSYRDLRGKGVVPYTOCKWEGELGTDLSVIPHGPNTVVRANIAAITE 120
DQ 93 HPFLHRYQRLSSYRDLRGKGVVPYTOCKWEGELGTDLSVIPHGPNTVVRANIAAITE 152
QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLQCGAGPPLNOS 180
DQ 153 SDKFFINGSNWEGILGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLQCGAGPPLNOS 212
QY 181 EVLASVSGSMIIGGDHSLYTGSLWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDK 240
DQ 213 EVLASVSGSMIIGGDHSLYTGSLWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDK 272
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKPEPDGFWLGEOLVCWQAGTTPNIPFPVIS 300
DQ 273 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKPEPDGFWLGEOLVCWQAGTTPNIPFPVIS 332
QY 301 LYLMEVNTQSFRTILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
DQ 333 LYLMEVNTQSFRTILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 392
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DQ 393 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 452
QY 421 CALFMLPLCLMVCQWRCRLRQOHDHDFADDISLLK 456
DQ 453 CALFMLPLCLMVCQWRCRLRQOHDHDFADDISLLK 488

RESULT 2
US-09-796-264-2
Sequence 2, Application US/09796264
Patent No. US2002049303A1
GENERAL INFORMATION:
APPLICANT: Tang, Jordan J.N.
APPLICANT: Lin, Xinli
APPLICANT: Koelsch, Gerald
TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
TITLE OF INVENTION: of Use Thereof
FILE REFERENCE: OMRF 179
CURRENT APPLICATION NUMBER: US/09/796,264
CURRENT FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 09/604,608
PRIOR FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/168,060
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: 60/177,836
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 60/178,368
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 60/210,292
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 488
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: Purified Memapsin 2
OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide
OTHER INFORMATION: residues
OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,
OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and
OTHER INFORMATION: 376-377 are residues in contact with the OM99-2
OTHER INFORMATION: inhibitor
OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,
OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and
OTHER INFORMATION: 220-224 are N-lobe Beta Strands
OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices
OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,
OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,
OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
OTHER INFORMATION: and 427-431 are C-lobe Helices
US-09-796-264-2

Query Match 100.0%; Score 2419; DB 10; Length 488;
Best Local Similarity 100.0%; Pred. No. 2.3e-224;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEPEGRGSEVEMVDNLGRKSGQGYVEMTSGSPQTLNLLVDTGSSNFAYGAAP 60
DQ 33 ETDEPEPEGRGSEVEMVDNLGRKSGQGYVEMTSGSPQTLNLLVDTGSSNFAYGAAP 92
QY 61 HPFLHRYQRLSSYRDLRGKGVVPYTOCKWEGELGTDLSVIPHGPNTVVRANIAAITE 120
DQ 93 HPFLHRYQRLSSYRDLRGKGVVPYTOCKWEGELGTDLSVIPHGPNTVVRANIAAITE 152
QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLQCGAGPPLNOS 180
DQ 153 SDKFFINGSNWEGILGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLQCGAGPPLNOS 212
QY 181 EVLASVSGSMIIGGDHSLYTGSLWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDK 240
DQ 213 EVLASVSGSMIIGGDHSLYTGSLWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDK 272
QY 241 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKPEPDGFWLGEOLVCWQAGTTPNIPFPVIS 300
DQ 273 SIVDSGTTNLRPKKVFEEAAVKSIIKAASSTKPEPDGFWLGEOLVCWQAGTTPNIPFPVIS 332
QY 301 LYLMEVNTQSFRTILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
DQ 333 LYLMEVNTQSFRTILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVV 392
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DQ 393 FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 452
QY 421 CALFMLPLCLMVCQWRCRLRQOHDHDFADDISLLK 456
DQ 453 CALFMLPLCLMVCQWRCRLRQOHDHDFADDISLLK 488

RESULT 3
US-09-845-226-2
Sequence 2, Application US/09845226
Patent No. US20020115600A1
GENERAL INFORMATION:
APPLICANT: Tang, Jordan J.N.
APPLICANT: Hong, Lin
APPLICANT: Ghosh, Arun K.
TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
FILE REFERENCE: OMRF 182
CURRENT APPLICATION NUMBER: US/09/845,226
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 09/603,713
PRIOR FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/168,060
PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: 60/177,836
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 60/178,368
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 60/210,292
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 488
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Purified Memapsin 2
OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide
OTHER INFORMATION: residues
OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,
OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and
OTHER INFORMATION: 376-377 are residues in contact with the OM99-2
OTHER INFORMATION: inhibitor
OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,
OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and
OTHER INFORMATION: 220-224 are N-lobe Beta Strands
OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices
OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,
OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,
OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
OTHER INFORMATION: and 427-431 are C-lobe Helices
US-09-843-226-2

Query Match 100.0%; Score 2419; DB 10; Length 488;
Best Local Similarity 100.0%; Pred. No. 2.3e-224;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEEPCRRGSGFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVGAAP 60
DB 33 ETDEEPEEPCRRGSGFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVGAAP 92
QY 61 HPFLHRYRQSLSTYRDLRKGVYVPYTOGKWEGLGTDLVSPHGPNTVTRANIAAITE 120
DB 93 HPFLHRYRQSLSTYRDLRKGVYVPYTOGKWEGLGTDLVSPHGPNTVTRANIAAITE 152
QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLQLCGAGFPLNOS 180
DB 153 SDKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLQLCGAGFPLNOS 212
QY 181 EVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 240
DB 213 EVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 272
QY 241 SIVDSGTTNLRPKPKVFEAAVKSIAASSTKFPDGFGLGEQLVCWQAGTTPNIPFVVIS 300
DB 273 SIVDSGTTNLRPKPKVFEAAVKSIAASSTKFPDGFGLGEQLVCWQAGTTPNIPFVVIS 332
QY 301 LYLMGEVNTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360
DB 333 LYLMGEVNTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 392
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DB 393 FDRARKRIGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 452
QY 421 CALFMLPLCLMWQWRCRLCRLRQHQHDDFADDISLLK 456
DB 453 CALFMLPLCLMWQWRCRLCRLRQHQHDDFADDISLLK 488

RESULT 4
US-09-795-903A-3
; Sequence 3, Application US/09795903A
; Patent No. US20020164760A1

GENERAL INFORMATION:
APPLICANT: Tang, Jordan J.N.
APPLICANT: Lin, Xinli
APPLICANT: Koelsch, Gerald
TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
TITLE OF INVENTION: Of Use Thereof
FILE REFERENCE: OMRF 179
CURRENT APPLICATION NUMBER: US/09/795,903A
CURRENT FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 09/604,608
PRIOR FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/168,060
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: 60/177,836
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 60/178,368
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 60/210,292
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 503
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: pro-memapsin 2
OTHER INFORMATION: Amino Acids 1-15 are vector-derived residues
OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide
OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter
OTHER INFORMATION: Amino Acids 16-456 are pro-memapsin 2-T1
OTHER INFORMATION: Amino Acids 16-421 are Promemapsin 2-T2
US-09-795-903A-3

Query Match 100.0%; Score 2419; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 2.4e-224;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEEPCRRGSGFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVGAAP 60
DB 48 ETDEEPEEPCRRGSGFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVGAAP 107
QY 61 HPFLHRYRQSLSTYRDLRKGVYVPYTOGKWEGLGTDLVSPHGPNTVTRANIAAITE 120
DB 108 HPFLHRYRQSLSTYRDLRKGVYVPYTOGKWEGLGTDLVSPHGPNTVTRANIAAITE 167
QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLQLCGAGFPLNOS 180
DB 168 SDKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLQLCGAGFPLNOS 227
QY 181 EVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 240
DB 228 EVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 287
QY 241 SIVDSGTTNLRPKPKVFEAAVKSIAASSTKFPDGFGLGEQLVCWQAGTTPNIPFVVIS 300
DB 288 SIVDSGTTNLRPKPKVFEAAVKSIAASSTKFPDGFGLGEQLVCWQAGTTPNIPFVVIS 347
QY 301 LYLMGEVNTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360
DB 348 LYLMGEVNTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 407
QY 361 FDRARKRIGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DB 408 FDRARKRIGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 467
QY 421 CALFMLPLCLMWQWRCRLCRLRQHQHDDFADDISLLK 456
DB 468 CALFMLPLCLMWQWRCRLCRLRQHQHDDFADDISLLK 503

RESULT 5
US-09-796-264-3

; Sequence 3, Application US/09796264
; Patent No. US20020049303A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
; TITLE OF INVENTION: of Use Thereof
; FILE REFERENCE: OMRF 179
; CURRENT APPLICATION NUMBER: US/097796,264
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/604,608
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Pro-memapsin 2
; OTHER INFORMATION: Amino Acids 1-15 are vector-derived residues
; OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide
; OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter
; OTHER INFORMATION: Amino Acids 16-456 are Pro-memapsin 2-T1
; OTHER INFORMATION: Amino Acids 16-421 are Promemapsin 2-T2
US-09-796-264-3

Query Match 100.0%; Score 2419; DB 10; Length 503;
Best Local Similarity 100.0%; Pred. No. 2.4e-224;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPGRGSGFVEMVDNLKSGGGYVYVMTVGSPPQTLNIIIVDTGSSNFVAGCAAP 60
DB 48 ETDEEPEEPGRGSGFVEMVDNLKSGGGYVYVMTVGSPPQTLNIIIVDTGSSNFVAGCAAP 107
QY 61 HPFLHRYQRLSSTYRDLRGVYVYPTQGWEGELGTLVSLPHGPNVTVRANIAAITE 120
DB 108 HPFLHRYQRLSSTYRDLRGVYVYPTQGWEGELGTLVSLPHGPNVTVRANIAAITE 167
QY 121 SSKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVPLNLSLQLCGAGFPLNQS 180
DB 168 SSKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVPLNLSLQLCGAGFPLNQS 227
QY 181 EVLASVGGSMIIGDHSLYTGLSWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 240
DB 228 EVLASVGGSMIIGDHSLYTGLSWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 287
QY 241 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVCVQAGTTPNNIPFV 300
DB 288 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVCVQAGTTPNNIPFV 347
QY 301 LYLMGEVTNOSFRITILPQQYLPRVEDVATSDDDCKYKFAISQSSTGTVMGAVIMEGFV 360
DB 348 LYLMGEVTNOSFRITILPQQYLPRVEDVATSDDDCKYKFAISQSSTGTVMGAVIMEGFV 407
QY 361 FDRARRKIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DB 408 FDRARRKIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 467
QY 421 CALFMPLPLCLMVCOWRCLRCLRQOHHDDFADDISLLK 456
DB 468 CALFMPLPLCLMVCOWRCLRCLRQOHHDDFADDISLLK 503

RESULT 6
US-09-845-226-3
; Sequence 3, Application US/09845226
; Patent No. US20020115600A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Hong, Lin
; APPLICANT: Ghosh, Arun K.
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
; FILE REFERENCE: OMRF 182
; CURRENT APPLICATION NUMBER: US/09/845,226
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 09/603,713
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Pro-memapsin 2
; OTHER INFORMATION: Amino Acids 1-15 are vector-derived residues
; OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide
; OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter
; OTHER INFORMATION: Amino Acids 16-456 are Pro-memapsin 2-T1
; OTHER INFORMATION: Amino Acids 16-421 are Promemapsin 2-T2
US-09-845-226-3

Query Match 100.0%; Score 2419; DB 10; Length 503;
Best Local Similarity 100.0%; Pred. No. 2.4e-224;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPGRGSGFVEMVDNLKSGGGYVYVMTVGSPPQTLNIIIVDTGSSNFVAGCAAP 60
DB 48 ETDEEPEEPGRGSGFVEMVDNLKSGGGYVYVMTVGSPPQTLNIIIVDTGSSNFVAGCAAP 107
QY 61 HPFLHRYQRLSSTYRDLRGVYVYPTQGWEGELGTLVSLPHGPNVTVRANIAAITE 120
DB 108 HPFLHRYQRLSSTYRDLRGVYVYPTQGWEGELGTLVSLPHGPNVTVRANIAAITE 167
QY 121 SSKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVPLNLSLQLCGAGFPLNQS 180
DB 168 SSKFFINGSNWEGILGLAYAEIARPDSDLPEFFDSLVKQTHVPLNLSLQLCGAGFPLNQS 227
QY 181 EVLASVGGSMIIGDHSLYTGLSWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 240
DB 228 EVLASVGGSMIIGDHSLYTGLSWYTPIRREWYVEIIVRVEINGODLKMDCKEYNDK 287
QY 241 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVCVQAGTTPNNIPFV 300
DB 288 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVCVQAGTTPNNIPFV 347
QY 301 LYLMGEVTNOSFRITILPQQYLPRVEDVATSDDDCKYKFAISQSSTGTVMGAVIMEGFV 360
DB 348 LYLMGEVTNOSFRITILPQQYLPRVEDVATSDDDCKYKFAISQSSTGTVMGAVIMEGFV 407
QY 361 FDRARRKIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
DB 408 FDRARRKIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 467
QY 421 CALFMPLPLCLMVCOWRCLRCLRQOHHDDFADDISLLK 456
DB 468 CALFMPLPLCLMVCOWRCLRCLRQOHHDDFADDISLLK 503

RESULT 7

US-09-794-927-4
: Sequence 4, Application US/09794927
: Patent No. US20010016324A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Bienkowski, Michael J.
: APPLICANT: Heinrichson, Robert L.
: APPLICANT: Parodi, Luis A.
: APPLICANT: Yan, Riqiang
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
: TITLE OF INVENTION: USES
: FILE REFERENCE: 28341/6280FG
: CURRENT APPLICATION NUMBER: US/09/794, 927
: PRIOR FILING DATE: 2001-02-27
: PRIOR APPLICATION NUMBER: 09/416, 901
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: 60/155, 493
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 09/404, 133
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: PCT/US99/20881
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 60/101, 594
: PRIOR FILING DATE: 1998-09-24
: NUMBER OF SEQ ID NOS: 73
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 501
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-794-927-4

Query Match 99.8%; Score 2414; DB 10; Length 501;
Best Local Similarity 99.8%; Pred. No. 7.le-224;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ETDEEPEPCRRGSGFVEMVDNLKSGQGYVVENTVGSPPOTLNILVDTGSSNFVCAAP	60
Db	46	ETDEEPEPCRRGSGFVEMVDNLKSGQGYVVENTVGSPPOTLNILVDTGSSNFVCAAP	105
Qy	61	HPFLHRYQRLSSTYRDLRKGYVVPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITE	120
Db	106	HPFLHRYQRLSSTYRDLRKGYVVPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITE	165
Qy	121	SDKFFINGSNWEGTLGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLHLCGAGFPLNQS	180
Db	166	SDKFFINGSNWEGTLGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLHLCGAGFPLNQS	225
Qy	181	EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNYDK	240
Db	226	EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNYDK	285
Qy	241	SIVDSGTTNLRPKKVEAAVKSIKAASSTKFPDGFWMGEQLVCWQAGTTPWNIFPVIS	300
Db	286	SIVDSGTTNLRPKKVEAAVKSIKAASSTKFPDGFWMGEQLVCWQAGTTPWNIFPVIS	345
Qy	301	LYLMGEVNTNSFRITILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV	360
Db	346	LYLMGEVNTNSFRITILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV	405
Qy	361	FDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLTMTIAYVMAAI	420
Db	406	FDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLTMTIAYVMAAI	465
Qy	421	CALFMPLCLMWCMWCRCLRQHQHDDFADDSLK 456	
Db	466	CALFMPLCLMWCMWCRCLRQHQHDDFADDSLK 501	

RESULT 8

US-09-795-847-4
: Sequence 4, Application US/09795847
: Patent No. US20010018208A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Bienkowski, Michael J.
: APPLICANT: Heinrichson, Robert L.
: APPLICANT: Parodi, Luis A.
: APPLICANT: Yan, Riqiang
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
: TITLE OF INVENTION: USES
: FILE REFERENCE: 28341/6280DE
: CURRENT APPLICATION NUMBER: US/09/795, 847
: PRIOR FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: 09/416, 901
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: 60/155, 493
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 09/404, 133
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: PCT/US99/20881
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 60/101, 594
: PRIOR FILING DATE: 1998-09-24
: NUMBER OF SEQ ID NOS: 73
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 501
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-795-847-4

Query Match 99.8%; Score 2414; DB 10; Length 501;
Best Local Similarity 99.8%; Pred. No. 7.le-224;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ETDEEPEPCRRGSGFVEMVDNLKSGQGYVVENTVGSPPOTLNILVDTGSSNFVCAAP	60
Db	46	ETDEEPEPCRRGSGFVEMVDNLKSGQGYVVENTVGSPPOTLNILVDTGSSNFVCAAP	105
Qy	61	HPFLHRYQRLSSTYRDLRKGYVVPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITE	120
Db	106	HPFLHRYQRLSSTYRDLRKGYVVPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITE	165
Qy	121	SDKFFINGSNWEGTLGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLHLCGAGFPLNQS	180
Db	166	SDKFFINGSNWEGTLGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLHLCGAGFPLNQS	225
Qy	181	EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNYDK	240
Db	226	EVLASVGGSMIIGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNYDK	285
Qy	241	SIVDSGTTNLRPKKVEAAVKSIKAASSTKFPDGFWMGEQLVCWQAGTTPWNIFPVIS	300
Db	286	SIVDSGTTNLRPKKVEAAVKSIKAASSTKFPDGFWMGEQLVCWQAGTTPWNIFPVIS	345
Qy	301	LYLMGEVNTNSFRITILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV	360
Db	346	LYLMGEVNTNSFRITILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV	405
Qy	361	FDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLTMTIAYVMAAI	420
Db	406	FDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLTMTIAYVMAAI	465
Qy	421	CALFMPLCLMWCMWCRCLRQHQHDDFADDSLK 456	
Db	466	CALFMPLCLMWCMWCRCLRQHQHDDFADDSLK 501	

RESULT 9
US-09-794-743-4
: Sequence 4, Application US/09794743

; Patent No. US20010021391A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND

; TITLE OF INVENTION: USES

; FILE REFERENCE: 28341/6280BC

; CURRENT APPLICATION NUMBER: US/09/794,743

; PRIOR FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 501

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-794-743-4

Query Match

Best Local Similarity

Matches 455; Conservative

0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGRGSGFYVEMVDNLKSGQGYVEMTGVSPQTLNLLVDTGSSNFVGAAP 60

DB 46 ETDEEPEEPRGRGSGFYVEMVDNLKSGQGYVEMTGVSPQTLNLLVDTGSSNFVGAAP 105

QY 61 HPFLHRYQRLSSTYRDLKRGVYVPTQGWEGELGTLVSIHPGPNVTVRANIAAITE 120

DB 106 HPFLHRYQRLSSTYRDLKRGVYVPTQGWEGELGTLVSIHPGPNVTVRANIAAITE 165

QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFDSLKVQTHVNPVLSLQLCGAGFPNQS 180

DB 166 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFDSLKVQTHVNPVLSLQLCGAGFPNQS 225

QY 181 EVLASVGSMTIIGGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNYDK 240

DB 226 EVLASVGSMTIIGGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNYDK 285

QY 241 SIIVDSGTTNLRPKKVFEEAAVKSIAKASSTKFPDGFGLGEQLVCWQAGTTPWNIFFPVIS 300

DB 286 SIIVDSGTTNLRPKKVFEEAAVKSIAKASSTKFPDGFGLGEQLVCWQAGTTPWNIFFPVIS 345

QY 301 LYLMEVNTNQSFRITLPOQYLRPEVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360

DB 346 LYLMEVNTNQSFRITLPOQYLRPEVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 405

QY 361 FDRARKRIGFAYSAHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420

DB 406 FDRARKRIGFAYSAHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465

QY 421 CALFMPLCLMVCQWRCLRCLRQHQHDDFADDSLLK 456

DB 466 CALFMPLCLMVCQWRCLRCLRQHQHDDFADDSLLK 501

RESULT 10

US-09-794-748-4

; Sequence 4, Application US/09794748

; Patent No. US20020037315A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND

; TITLE OF INVENTION: USES

; FILE REFERENCE: 28341/6280JL

; CURRENT APPLICATION NUMBER: US/09/794,748

; PRIOR FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 501

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-794-748-4

Query Match

Best Local Similarity

Matches 455; Conservative

0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGRGSGFYVEMVDNLKSGQGYVEMTGVSPQTLNLLVDTGSSNFVGAAP 60

DB 46 ETDEEPEEPRGRGSGFYVEMVDNLKSGQGYVEMTGVSPQTLNLLVDTGSSNFVGAAP 105

QY 61 HPFLHRYQRLSSTYRDLKRGVYVPTQGWEGELGTLVSIHPGPNVTVRANIAAITE 120

DB 106 HPFLHRYQRLSSTYRDLKRGVYVPTQGWEGELGTLVSIHPGPNVTVRANIAAITE 165

QY 121 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFDSLKVQTHVNPVLSLQLCGAGFPNQS 180

DB 166 SDKFFINGSNWEGILGLAYAEIARPDSDLEPFDSLKVQTHVNPVLSLQLCGAGFPNQS 225

QY 181 EVLASVGSMTIIGGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNYDK 240

DB 226 EVLASVGSMTIIGGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNYDK 285

QY 241 SIIVDSGTTNLRPKKVFEEAAVKSIAKASSTKFPDGFGLGEQLVCWQAGTTPWNIFFPVIS 300

DB 286 SIIVDSGTTNLRPKKVFEEAAVKSIAKASSTKFPDGFGLGEQLVCWQAGTTPWNIFFPVIS 345

QY 301 LYLMEVNTNQSFRITLPOQYLRPEVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360

DB 346 LYLMEVNTNQSFRITLPOQYLRPEVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 405

QY 361 FDRARKRIGFAYSAHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420

DB 406 FDRARKRIGFAYSAHVHDEFTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 465

QY 421 CALFMPLCLMVCQWRCLRCLRQHQHDDFADDSLLK 456

DB 466 CALFMPLCLMVCQWRCLRCLRQHQHDDFADDSLLK 501

RESULT 11

US-09-794-925-4

; Sequence 4, Application US/09794925

; Patent No. US20020064819A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

APPLICANT: Heinrichson, Robert L.
APPLICANT: Parodi, Luis A.
APPLICANT: Yan, Riqiang
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
FILE REFERENCE: 28341/6280H1
CURRENT APPLICATION NUMBER: US/09/794,925
CURRENT FILING DATE: 2001-02-27
PRIOR APPLICATION NUMBER: 09/416,901
PRIOR FILING DATE: 1999-10-13
PRIOR APPLICATION NUMBER: 60/155,493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 09/404,133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/1101,594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 501
TYPE: PRT
ORGANISM: Homo sapiens
US-09-794-925-4

Query Match 99.8%; Score 2414; DB 10; Length 501;
Best Local Similarity 99.8%; Pred. No. 7.1e-224;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ETDEEPEEPPGRRGGSFVEMVNDLGRKSGQGYVYVEMTVGSPQTLNLLVDTGSSNFAVGAAP 60
Db 46 ETDEEPEEPPGRRGGSFVEMVNDLGRKSGQGYVYVEMTVGSPQTLNLLVDTGSSNFAVGAAP 105

Qy 61 HPFLHRYQRLSTYRDLRKGYVVPYTOGKWEGLGTDLVSPHGPNTVVRANTAATE 120
Db 106 HPFLHRYQRLSTYRDLRKGYVVPYTOGKWEGLGTDLVSPHGPNTVVRANTAATE 165

Qy 121 SDKFFINGSNWEGTLGAYAEIARPDSDLPPFDSLVKQTHVFNLSLQLCGAGFPLNOS 180
Db 166 SDKFFINGSNWEGTLGAYAEIARPDSDLPPFDSLVKQTHVFNLSLQLCGAGFPLNOS 225

Qy 181 EVLASVGSMTIGDHSHTYGLSWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 240
Db 226 EVLASVGSMTIGDHSHTYGLSWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 285

Qy 241 SIVDSGTTNLRPKKVEAAVKSIAASSTKFPDGFGLGEQLVCHQAGTTTWNIFPVIS 300
Db 286 SIVDSGTTNLRPKKVEAAVKSIAASSTKFPDGFGLGEQLVCHQAGTTTWNIFPVIS 345

Qy 301 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 346 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 405

Qy 361 FDRARKRIGFAVSACHVHDEFTAAVEGPPVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 420
Db 406 FDRARKRIGFAVSACHVHDEFTAAVEGPPVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 465

Qy 421 CALFMLPLCLMVCMQWRCRLCRLQOHHDDFADDISLLK 456
Db 466 CALFMLPLCLMVCMQWRCRLCRLQOHHDDFADDISLLK 501

RESULT 12
US-09-681-442-4
Sequence 4, Application US/09681442
Patent No. US20020081634A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Blenkowski, Michael J.
APPLICANT: Heinrichson, Robert L.
APPLICANT: Parodi, Luis A.
APPLICANT: Yan, Riqiang

TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
FILE REFERENCE: 28341/6280FG
CURRENT APPLICATION NUMBER: US/09/681,442
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 09/416,901
PRIOR FILING DATE: 1999-10-13
PRIOR APPLICATION NUMBER: 60/155,493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 09/404,133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: PCT/US99/20881
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/1101,594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 501
TYPE: PRT
ORGANISM: Homo sapiens
US-09-681-442-4

Query Match 99.8%; Score 2414; DB 10; Length 501;
Best Local Similarity 99.8%; Pred. No. 7.1e-224;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ETDEEPEEPPGRRGGSFVEMVNDLGRKSGQGYVYVEMTVGSPQTLNLLVDTGSSNFAVGAAP 60
Db 46 ETDEEPEEPPGRRGGSFVEMVNDLGRKSGQGYVYVEMTVGSPQTLNLLVDTGSSNFAVGAAP 105

Qy 61 HPFLHRYQRLSTYRDLRKGYVVPYTOGKWEGLGTDLVSPHGPNTVVRANTAATE 120
Db 106 HPFLHRYQRLSTYRDLRKGYVVPYTOGKWEGLGTDLVSPHGPNTVVRANTAATE 165

Qy 121 SDKFFINGSNWEGTLGAYAEIARPDSDLPPFDSLVKQTHVFNLSLQLCGAGFPLNOS 180
Db 166 SDKFFINGSNWEGTLGAYAEIARPDSDLPPFDSLVKQTHVFNLSLQLCGAGFPLNOS 225

Qy 181 EVLASVGSMTIGDHSHTYGLSWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 240
Db 226 EVLASVGSMTIGDHSHTYGLSWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNDK 285

Qy 241 SIVDSGTTNLRPKKVEAAVKSIAASSTKFPDGFGLGEQLVCHQAGTTTWNIFPVIS 300
Db 286 SIVDSGTTNLRPKKVEAAVKSIAASSTKFPDGFGLGEQLVCHQAGTTTWNIFPVIS 345

Qy 301 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 346 LYLMEVTNQSFRITILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYV 405

Qy 361 FDRARKRIGFAVSACHVHDEFTAAVEGPPVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 420
Db 406 FDRARKRIGFAVSACHVHDEFTAAVEGPPVTLDMEDCGYNIPQTDDESTLMTIAYVMAAI 465

Qy 421 CALFMLPLCLMVCMQWRCRLCRLQOHHDDFADDISLLK 456
Db 466 CALFMLPLCLMVCMQWRCRLCRLQOHHDDFADDISLLK 501

RESULT 13
US-09-969-671A-2
Sequence 2, Application US/09969671A
Publication No. US20030036112A1
GENERAL INFORMATION:
APPLICANT: CHAPMAN, CONRAD G.
APPLICANT: MURPHY, KAY
APPLICANT: POWELL, DAVID J.
APPLICANT: SMITH, TRUDI S.
TITLE OF INVENTION: ASP2
FILE REFERENCE: GH-70368-D1
CURRENT APPLICATION NUMBER: US/09/969,671A
CURRENT FILING DATE: 2001-10-03

```
; PRIORITY NUMBER: UK 9701684.4
; PRIOR FILING DATE: 1997-01-28
; PRIOR APPLICATION NUMBER: 09/009,191
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 09/694,200
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-969-671A-2
```

Query Match	99.98;	Score	2413;	DB	9;	Length	501;
Best Local Similarity	99.98;	Pred.	No. 8.9e-224;				
Matches	455;	Conservative	0;	Mismatches	1;	Indels	0;
Qy	1	ETDEPEPEGRGRGSFVEMVDNLRC	KGQGYVEMTVGSP	PQTNLILVD	TGSSNFAVGAAP	60	
Db	46	ETDEPEPEGRGRGSFVEMVDNLRC	KGQGYVEMTVGSP	PQTNLILVD	TGSSNFAVGAAP	105	
Qy	61	HPFLHRYQRLSTYRDLRKGVVYPT	YQGWKEGELG	TDLVISPHGPNVT	VRANIAAITE	120	
Db	106	HPFLHRYQRLSTYRDLRKGVPEPY	TQGWKEGELG	TDLVISPHGPNVT	VRANIAAITE	165	
Qy	121	SDKFFINGSNWEGLGLAYAEIAR	PDPSLEPF	FDLSVKQTHV	PNLFSLQLCGAGFPLNQS	180	
Db	166	SDKFFINGSNWEGLGLAYAEIAR	PDPSLEPF	FDLSVKQTHV	PNLFSLQLCGAGFPLNQS	225	
Qy	181	EVLASVGSMITIGIDHS	LYTGS	LWYTPIRREWY	EVIIVRVEINGO	LKMDCKREYN	240
Db	226	EVLASVGSMITIGIDHS	LYTGS	LWYTPIRREWY	EVIIVRVEINGO	LKMDCKREYN	285
Qy	241	SIVDSGTTNLR	LPKKVF	EAAVKSKAAS	STKEKFPD	GWLGQLCWQAGTTP	300
Db	286	SIVDSGTTNLR	LPKKVF	EAAVKSKAAS	STKEKFPD	GWLGQLCWQAGTTP	345
Qy	301	LYLGMGEVNTGS	FRITLIPQOY	LRPVEDVAT	SQDDCYKFAISQS	STGTVMGAVIMEG	360
Db	346	LYLGMGEVNTGS	FRITLIPQOY	LRPVEDVAT	SQDDCYKFAISQS	STGTVMGAVIMEG	405
Qy	361	FDRAKRRIGFAVS	ACHVHDEF	RTAAVGGP	FVTLDMEDCGY	NIPOTDESTLMTIAYVMAAI	420
Db	406	FDRAKRRIGFAVS	ACHVHDEF	RTAAVGGP	FVTLDMEDCGY	NIPOTDESTLMTIAYVMAAI	465
Qy	421	CALFMFLPLCLMVC	OWRCLRLC	RQOHDDF	FADDISLKK	456	
Db	466	CALFMFLPLCLMVC	OWRCLRLC	RQOHDDF	FADDISLKK	501	

RESULT 14
US-09-794-927-8
Sequence 8, Application US/09794927
Patent NO. US20010016324A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Bienkowski, Michael J.
APPLICANT: Henrikson, Robert L.
APPLICANT: Parodi, Luis A.
APPLICANT: Yan, Riqiang
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
TITLE OF INVENTION: USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 28341/6280FG
CURRENT APPLICATION NUMBER: US/09/794,927
CURRENT FILING DATE: 2001-02-27
PRIORITY APPLICATION NUMBER: 05/416,901
PRIORITY FILING DATE: 1999-10-13
PRIORITY APPLICATION NUMBER: 60/155,493
PRIORITY FILING DATE: 1999-09-23
PRIORITY APPLICATION NUMBER: 09/404,133
PRIORITY FILING DATE: 1999-09-23

```

; PRIORITY APPLICATION NUMBER: PCT/US99/20881
; PRIORITY FILING DATE: 1999-09-23
; PRIORITY APPLICATION NUMBER: 60/101,594
; PRIORITY FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 501
; TYPE: prt
; ORGANISM: Mus musculus
US-09-794-927-8

```

Query Match	98.98%	Score 2390;	DB 10;	Length 501;
Best Local Similarity	98.2%	Pred. No. 1.4e-221;		
Matches 448; Conservative 4;	Mismatches 4;	Indels 0;	Caps 0;	
Qy 1	ETDEEPEPCRRGSFVEMVDNLRCKSGQYGVVEMTVGSPQTLLNLVDTGSSNFVAGCAAP	60		
Db				
Qy 46	ETDEESEPGRGRGSFVEMVDNLRCKSGQYGVVEMTVGSPQTLLNLVDTGSSNFVAGCAAP	105		
Db				
Qy 61	HPFLHRYQRQLSSTYRDLRKGVVVPYVYTGQKWEGELGTLVSTPHGPNVTVRANIAAITE	120		
Db				
Qy 106	HPFLHRYQRQLSSTYRDLRKGVVVPYVYTGQKWEGELGTLVSTPHGPNVTVRANIAAITE	165		
Db				
Qy 121	SDKFFINGSNWEGITGLGAYAEIARPDSDLSPFFDSLKVQTHVPNLFSJQLCGAGFPLNQSL	180		
Db				
Qy 166	SDKFFINGSNWEGITGLGAYAEIARPDSDLSPFFDSLKVQTHVPNLFSJQLCGAGFPLNOT	225		
Db				
Qy 181	EVLASVGSMSIIGGDHSLTYTSLWYTPIRREWYVEIIVRVEINGODLKMDCKEYNVDK	240		
Db				
Qy 226	EALASVGSMSIIGGDHSLTYTSLWYTPIRREWYVEIIVRVEINGODLKMDCKEYNVDK	285		
Db				
Qy 241	SIVDSGTTLNLRLPKKVEAAVKSLKAASSTKEKPDGFMWLGEOVCWQAGTTPWNIFPVIS	300		
Db				
Qy 286	SIVDSGTTLNLRLPKKVEAAVKSLKAASSTKEKPDGFMWLGEOVCWQAGTTPWNIFPVIS	345		
Db				
Qy 301	LYLMGEVYNTGSFRITILUPOOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVV	360		
Db				
Qy 346	LYLMGEVYNTGSFRITILUPOOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVV	405		
Db				
Qy 361	FDRAARKRIGFAVSACHVHDEFRTAAVSGPFTLDMEDCGYNIPTDDESTLMTIAYVMAAI	420		
Db				
Qy 406	FDRAARKRIGFAVSACHVHDEFRTAAVSGPFTLDMEDCGYNIPTDDESTLMTIAYVMAAI	465		
Db				
Qy 421	CALFMPLCLCUMVCOWRCLRCRQOHDDFADDDISLLK	456		
Db				
Qy 466	CALFMPLCLCUMVCOWRCLRCRQOHDDFADDDISLLK	501		
Db				

```

RESULT 15
US-09-795-847-8
: Sequence 8, Application US/09795847
: Patent No. US20010018208A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Bienkowski, Michael J.
: APPLICANT: Heinrikson, Robert L.
: APPLICANT: Parodi, Luis A.
: APPLICANT: Yan, Riqiang
:
: TITLE OF INVENTION: ALZHEIMER'S DISEASE
:
: TITLE OF INVENTION: USES
:
: FILE REFERENCE: 28341/6280DE
: CURRENT APPLICATION NUMBER: US/09/795,847
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: 09/416,901
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: 60/155,493
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 09/404,133
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: PCT/US99/20881
: PRIOR FILING DATE: 1999-09-23

```

; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-795-847-8

Query Match	98.8%	Score 2390;	DB 10;	Length 501;
Best local Similarity	98.2%	Pred. No. 1.4e-221;		
Matches 448;	Conservative 4;	Mismatches 4;	Indels 0;	Gaps 0;

Qy	1	ETDEEPEPGRRGSFVEMVNLKSGQGYVEMTVGSPPTLNILVDTGSSNFAVGAAP	60
Db	46	ETDEEPEPGRRGSFVEMVNLKSGQGYVEMTVGSPPTLNILVDTGSSNFAVGAAP	105
Qy	61	HPFLHRYVQRLSSTYRDLRGVYVPTQKGEGELGTDLVSIPIHGPNTVVRANIAAITE	120
Db	106	HPFLHRYVQRLSSTYRDLRGVYVPTQKGEGELGTDLVSIPIHGPNTVVRANIAAITE	165
Qy	121	SDKFFINGSNWEGILGLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLQLCGAGFPLNQ	180
Db	166	SDKFFINGSNWEGILGLAYAEIARPDSDLPPFDLSLVKQTHIPNIFSLQLCGAGFPLNQ	225
Qy	181	EVLASVGSMTIIGDHSILVTSGLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNYDK	240
Db	226	EALASVGSMTIIGDHSILVTSGLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNYDK	285
Qy	241	SIVDSGTTNLRPKKVEAAVKSTKAASSTKPPDGFMLGELVCMQAGTTPWNIFPVIS	300
Db	286	SIVDSGTTNLRPKKVEAAVKSTKAASSTKPPDGFMLGELVCMQAGTTPWNIFPVIS	345
Qy	301	LYLMGEVNTQSFRTILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV	360
Db	346	LYLMGEVNTQSFRTILPQOYLRPVEDVATSQDDCYKFAVQSSTGTVMGAVIMEGFYV	405
Qy	361	FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAI	420
Db	406	FDRARKRIGFAVSACHVHDEFRTAAVEGPFVTAADMEDCGYNIPOTDESTLMTIAYVMAAI	465
Qy	421	CALFMLPLCLMVQWRCRLRQOHHDDFADDISLLK	456
Db	466	CALFMLPLCLMVQWRCRLRQOHHDDFADDISLLK	501

Search completed: March 4, 2003, 10:08:31
Job time : 18 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:00:56 ; Search time 21 Seconds
(without alignments)
2087.490 Million cell updates/sec

Title: US-09-723-722A-43
Perfect score: 2419
Sequence: 1 ETDEPEPEPGRGSGFVEMVD.....CLRLRQHQHDFADDISLLK 456
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2419	100.0	501	2 A59090	aspartic proteinas
2	308	12.7	384	2 JC7574	pepsinogen A - Afr
3	308	12.7	385	2 JC7575	pepsinogen A - bul
4	306	12.6	387	2 JC7573	pepsin (EC 3.4.23.
5	303	12.5	383	2 JC7573	pepsinogen C - Afr
6	302.5	12.5	388	1 S19682	pepsin A (EC 3.4.2
7	301	12.4	382	1 PECH	pepsin A (EC 3.4.2
8	299.5	12.4	396	2 A34401	cathepsin E (EC 3.
9	298.5	12.3	383	2 A41443	pepsin (EC 3.4.23.
10	298.5	12.3	384	2 A39314	gastricsin (EC 3.4
11	296	12.2	387	2 C38302	pepsin (EC 3.4.23.
12	296	12.2	391	2 A43356	cathepsin E (EC 3.
13	295.5	12.2	412	1 KHHUD	cathepsin D (EC 3.
14	295	12.2	387	2 C38302	pepsin (EC 3.4.23.
15	289.5	12.0	388	1 S19684	pepsin A (EC 3.4.2
16	287.5	11.9	444	2 T24204	hypothetical prote
17	287	11.9	407	1 KHR7D	cathepsin D (EC 3.
18	286	11.8	387	2 E38302	pepsin (EC 3.4.23.
19	285.5	11.8	398	2 S66465	cathepsin E (EC 3.
20	285	11.8	398	2 I51185	cathepsin D (EC 3.
21	283.5	11.7	388	1 PEMAAR	pepsin A (EC 3.4.2
22	282.5	11.7	388	1 PEHU	pepsin A (EC 3.4.2
23	281.5	11.6	381	1 CMSHB	chymosin (EC 3.4.2
24	281.5	11.6	388	1 PEMAQJ	pepsin A (EC 3.4.2
25	281.5	11.6	388	2 A30142	pepsin A (EC 3.4.2
26	279.5	11.6	388	2 B30142	pepsin A (EC 3.4.2
27	279.5	11.6	410	1 KMSD	cathepsin D (EC 3.
28	278.5	11.5	386	1 PEPG	pepsin A (EC 3.4.2
29	278	11.5	387	2 JC7245	pepsinogen A - com

30	277.5	11.5	380	2 I47176	chymosin (EC 3.4.2
31	277.5	11.5	396	2 J36865	cathepsin E (EC 3.
32	276	11.4	389	2 JF0371	pepsin C (EC 3.4.2
33	273.5	11.3	381	1 CMBO	chymosin (EC 3.4.2
34	270.5	11.2	377	1 PEMOCJ	gastricsin (EC 3.4
35	270.5	11.2	389	2 A38302	pepsin (EC 3.4.23.
36	270	11.2	376	2 I45856	aspartic proteinas
37	268.5	11.1	344	1 KHPGD	cathepsin D (EC 3.
38	267.5	11.1	381	2 JC7247	prochymosin - comm
39	266	11.0	380	2 S03433	canidapepsin (EC
40	266	11.0	405	2 A25379	saccharopepsin (EC
41	264	10.9	396	2 T47207	aspartic proteinas
42	263.5	10.9	388	2 JC7246	pepsinogen C - com
43	262.5	10.9	394	2 B43356	gastricsin (EC 3.4
44	261.5	10.8	387	2 A45117	aspartic proteinas
45	261.5	10.8	388	2 A29937	gastricsin (EC 3.4

ALIGNMENTS

RESULT 1
A59090
aspartic proteinase (EC 3.4.23.-) BACE precursor - human
N;Alternate names: beta-secretase; beta-site APP cleaving enzyme
C;Species: Homo sapiens (man)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000
C;Accession: A59090
M;Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Teplio
M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro
Science 286, 735-741, 1999
A;Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran
A;Reference number: A59090; MUID:20002972; PMID:10531052
A;Note: submitted to GenBank, September 1999
A;Accession: A59090
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-501 <VAS>
A;Cross-references: GB:AF190725; NID:g6118538; PIDN:AAF04142.1; PID:g6118539
C;Genetics:
A;Gene: BACE
C;Superfamily: beta-secretase
C;Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein; hydrolase;
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-45/Domain: propeptide #status predicted <PRO>
F;46-501/Product: acid proteinase BACE #status predicted <MAT>
F;461-477/Domain: transmembrane #status predicted <TRN>
F;93,289/Active site: Asp #status predicted
F;153,172,223,354/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;330-380/Disulfide bonds: #status predicted

Query Match 100.0%; Score 2419; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 3.8e-192;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ETDEPEPEPGRGSGFVEMVDNLKSGQGYVEMTVGSPPTNLIVDTGSSNFVAGAAP	60
Db	46	ETDEPEPEPGRGSGFVEMVDNLKSGQGYVEMTVGSPPTNLIVDTGSSNFVAGAAP	105
Qy	61	HPFHRYRQRLSSTYRDLRGVYVPTVGKWEGLGDLVSIPIHPGNVTVRANIAATE	120
Db	106	HPFHRYRQRLSSTYRDLRGVYVPTVGKWEGLGDLVSIPIHPGNVTVRANIAATE	165
Qy	121	SDRFFINGSNWEGILGLAYAEIARPDPSLEFFDLSLVKQTHVPSLQLCAGFPLNQS	180
Db	166	SDKFFINGSNWEGILGLAYAEIARPDPSLEFFDLSLVKQTHVPSLQLCAGFPLNQS	225
Qy	181	EVLASVGSGMIIGIDHSLSYTGSLWYTPIRREWYVEYIIVRVEINGDLKMDCKEYNDK	240
Db	226	EVLASVGSGMIIGIDHSLSYTGSLWYTPIRREWYVEYIIVRVEINGDLKMDCKEYNDK	285
Qy	241	SIVDSGTNLRPKKVFEEAAVKSKAASSTKFFDGFGLGQVLCVQAGTTPWNIFFVIS	300

Db	286	STVDSGTTNLRLPKKVFEAAVKSIAASSTKFPDGFGLBQLCYCQAGTTPNIFPVIS	345
Qy	301	LYLMGEVNTNQSFRTITLPOQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV	360
Db	346	LYLMGEVNTNQSFRTITLPOQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYV	405
Qy	361	FDRARKRIGFAVSACHVHDEFFRTAAVEGPEVTTLDMEDCGYNIPQTDSTLMTIAYVMAAI	420
Db	406	FDRARKRIGFAVSACHVHDEFFRTAAVEGPEVTTLDMEDCGYNIPQTDSTLMTIAYVMAAI	465
Qy	421	CALFWMPLCLMVCQWRCLRCLRQOHDHFADDISLLK	456
Db	466	CALFWMPLCLMVCQWRCLRCLRQOHDHFADDISLLK	501

RESULT. 2

JC7574
 pepsinogen A - African clawed frog
 C:Species: *Xenopus laevis* (African clawed frog)
 C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
 C:Accession: JC7574; PC7119
 R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
 J. Biochem. 129, 147-153, 2001
 A:Title: Amphibian pepsinogens: Purification and characterization of *Xenopus pepsinogens*
 A:Reference number: JC7573; MUID:21064922; PMID:11134969
 A:Contents: Stomach
 A:Accession: JC7574
 A:Molecule type: mRNA
 A:Residues: 1-384 <IKU>
 A:Cross-references: DDBJ:AB045380
 A:Accession: PC7119
 A:Molecule type: protein
 A:Residues: 16-35;57-76 <IK2>
 C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like a
 C:Genetics:
 A:Gene: Pga
 C:Superfamily: pepsin
 C:Keywords: stomach; zymogen

RESULT 3

pepsinogen A - bullfrog
JC7375
C:Species: Rana catesbeiana (bullfrog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001

C:Accession: JC7575
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and Characterization of xenopus pepsinogen
A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7575
A:Molecule type: mRNA
A:Residues: 1-385 <IKU>
A:Cross-references: DDBJ:AB045376
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsinogen
C:Genetics:
A:Gene: PgA
C:Superfamily: pepsin
C:Keywords: stomach; zymogen

DESCRIPTION

B38302
 pepsin (EC 3.4.23.-) II-1 precursor - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 14-Jun-1991 #sequence_revision 20-Sep-1991 #text_change 23-Feb-1997
 C:Accession: B38302
 R:Kageyama, T.; Tanabe, K.; Koiwai, O.
 J. Biol. Chem. 265, 17031-17038, 1990
 A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens
 A:Reference number: A38302; M01d:91009127; PMID:2129556
 A:Accession: B38302
 A:Status: preliminary
 A:Molecule Type: mRNA
 A:Residues: 1-387 <KAG>
 A:Cross-references: GB:M59235; GB:J05638
 C:Superfamily: pepsin
 C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

C:Species: Rana catesbeiana (bullfrog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001

Db 135 TYTGSMTCGILGYDVTVKV---GNIEDTNOIFGLSKTEPGITFLV--APPDGLGLAYPSI 189
QY 143 ARPDSLEPFDSLVKQTHV-PNLFSLQLCGAGFPLNQSEVLASVGGSMIGIDHSLYT 201
Db 190 SASDAT--PVFDNMWNEGLVSEDLFSVYLSNG-----EKGSMWFGIDSSYIT 237
QY 202 GSLWTPPIRREWYEVIIIVRVEINGODLKM--DCKEYNYDKSIVDSGTTNLRPKKVFPA 259
Db 238 GSLNWVPVSHGYWQITMDSITINGETIACADSC-----QAVVDGTGTSLLAGPTSAISK 291
QY 260 AVKSKAASSTKFPDGFGLWGLQV-CWQAGTTPWNIFFVISLYLMGEVTSNFRITILP 318
Db 292 IQSYICASKNL-----LGENIITSCSAIDSLPDIVF-----TINN 325
QY 319 QOYLPRVED-VATSDDC---YKFAISOSSTGT--VMGAVIMEGFVYVDFDRARKRIGFAV 372
Db 326 VOYPLPASAYILKEDDCLSGFDGGMNLDTSYCELWILGDVFTROYTTFVDRANNQVGLAA 385
QY 373 SA 374
Db 386 AA 387

RESULT 5
JC7573
pepsinogen C - African clawed frog
N:Alternate names: progastricin
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
A:Accession: JC7573; PC7118
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens
A:Reference number: JC7573; PMID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7573
A:Molecule type: mRNA
A:Residues: 1-383 <IKU>
A:Cross-references: DDBJ:AB045379
A:Accession: PC7118
A:Molecule type: protein
A:Residues: 17-68 <IK2>
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like a
C:Genetics:
A:Gene: PgC
C:Superfamily: pepsin
C:Keywords: stomach; zymogen

Query Match 12.5%; Score 303; DB 2; Length 383;
Best Local Similarity 27.4%; Pred. No. 3.4e-17;
Matches 104; Conservative 57; Mismatches 122; Indels 96; Gaps 17;

QY 30 YVVEVTGSPPTNLTLVDTGSSNFAVGA-----APHPFLHRYQRLSTYRDLRK 81
Db 67 YGETSIGTPPNFLVDTGSSNLWVASTYCSQACTNHPL----FNPSQSSTYSNQ 122
QY 82 GVVPVPTQCKWCELCGTDLVSLPHGPNVTVRANIAITSDKFFINGSN----WEGILG 136
Db 123 QFSLQVGTGSLGILGYDVTI---ONVALSQEFGSLSETP----GTFNVYAQFQILG 175
QY 137 LAYAEIARPDSDSLEPPFDSLVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGID 196
Db 176 LAYPSIA--VGGATTVMQGMQO---NLLNQPI--FGYLSQS--SQNGGEVAFGGVD 225
QY 197 HSLYTGSLWYTPPIRREWYEVIIIVRVEINGOD---LKMDCKEYNYDKSIVDSGTTNLRPL 253
Db 226 QNYVTGQIYVPTSETYQWIGIQGFSINGQATGWCSSQGC-----QAVVDGTGTSLLTAP 279
QY 254 KKVFEAAVKSIAASSTKFPDGFGLWGLQV-CWQAGTTPWNI--FPVISLYLMG----- 305
Db 280 QSVFSLIQSIGAQQDN-----GOYVWCS-----NQNLPTISFTISGSFPLP 325
QY 306 ---EVTNQS-----FRITILPQOYLPRVEDVATSDDCYKFAISQSSTGTVMGAVIME 355

Db 326 PSAYVLQSSGYCTIGIMPTILPSONQPL-----WTLGDFLFR 364
QY 356 GEYVVFDRARKRIGFAVSA 374
Db 365 EYISVYDLGNNOVGFAATA 383

RESULT 6
S19682
pepsin A (EC 3.4.23.1) 4 precursor - Japanese macaque
N:Alternate names: pepsinogen A isozyme 4
C:Species: Macaca fuscata (Japanese macaque)
C:Date: 22-Nov-1993 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999
A:Accession: S19682; S16065
R:Kageyama, T.; Tanabe, K.; Koiwai, O.
Eur. J. Biochem. 202, 205-215, 1991
A:Title: Development-dependent expression of isozymogens of monkey pepsinogens and st
A:Reference number: S19681; PMID:92037645; PMID:1935977
A:Accession: S19682
A:Molecule type: mRNA
A:Residues: 1-388 <KAG>
A:Cross-references: EMBL:X59753; NID:g38070; PIDN:CAA2425.1; PID:g38071
A:Note: parts of sequence, including amino ends of pepsinogen and activation intermed
C:Comment: This is a minor component of pepsin at all post-partum stages.
C:Superfamily: pepsin
C:Comment: Although two-step activation is observed, activation is predominantly a o
C:Keywords: aspartic proteinase; gastric juice; hydrolase; phosphoprotein; protein di
F:1-15/Domain: signal sequence #status predicted <Sig>
F:16-388/Product: pepsinogen A 4 #status experimental <PPT>
F:16-62/Domain: activation peptide #status experimental <APT>
F:63-388/Product: pepsin A 4 #status experimental <ENZ>
F:38-39/cleavage site: Leu-Lys (pepsin) #status experimental
F:62-63/cleavage site: Leu-Ile (pepsin) #status experimental
F:94,277/Active site: Asp #status predicted
F:107-112,268-272,311-344/Disulfide bonds: #status predicted
F:130/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 12.5%; Score 302.5; DB 1; Length 388;
Best Local Similarity 27.6%; Pred. No. 3.8e-17;
Matches 107; Conservative 64; Mismatches 133; Indels 83; Gaps 17;

QY 3 DEEPEEPGRKGSFEMVDNLGRKSGGYVYVTVGSPPTNLTLVDTGSSNFAVGAAPHP 62
Db 64 DEQPLE-----NYLDV-----EYFGTIGTGAQNFVTVFDTGSSNLWV---PSV 105
QY 63 FL-----HRYQRLSTYRDLRKGVVYPTQGWEGELGTDLVSLPHGPNVTVRANI 115
Db 106 YCYSLACMDHNLFPQDSSYRATSKVTSITYGTGSMTGILGYDVTVKV---GGISDTNQI 162
QY 116 AATTESDK-PFINGSNWEGTLGLAYAEIARPDSDSLEPPFDSLVKQTHV-PNLFSLQLCGA 173
Db 163 FGLSETPGEFFLFAPDGLGLAYPSIS---SSGATPVFDNIWNQRLVSDLSFVLSAD 220
QY 174 GFLPNQSEVLASVGGSMIIGGIDHSLYTPIRREWYEVIIIVRVEINGODL--KM 231
Db 221 ---DQS-----GSVVFGIDSSYITGSLNWPVSVVEGYWQISVDSITMCKGTACAK 270
QY 232 DCKEYNYDKSIVDSGTTNLRPKKVFPAVAAVKSIAASSTKFPDGFGLWGLQV-CWQAGT 290
Db 271 GC-----QAVVDGTGTSLLTGTPTSPIANIQSDIGASENSD-----GEMVYSCSAISS 316
QY 291 TPNWIFPVLSLYLMGEVTSNFRITILPQOY-LRPVEDVATSDDCYK-----FAISQSS 344
Db 317 LPDIVF-----TINGVOYPLPPSAYILQSOGSCSTSGFGQMDVPTFESG 358
QY 345 TGTVMGAVIMEGFVYVDFDRARKRIGFA 371
Db 359 ELWILGDVFTROYTTFVDRANNQVGLA 385

RESULT 7
PECH

pepsin A (EC 3.4.23.1) precursor - chicken
N:Alternate names: pepsinogen A
C:Species: Gallus gallus (chicken)
C:Date: 18-Apr-1984 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C:Accession: J0370; A00984
R:Sakamoto, N.; Saiga, H.; Yasugi, S.
Biochem. Biophys. Res. Commun. 250, 420-424, 1998
A:Title: Analysis of temporal expression pattern and cis-regulatory sequences of chicken
A:Reference number: JE0370; MUID:98440813; PMID:9753645
A:Accession: JE0370
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-382 <SA>
A:Cross-references: GB:AB025281; NID:q4589837; PIDN:BAA76891.1; PID:g4589838
R:Baudys, M.; Kostka, V.
Eur. J. Biochem. 136, 89-99, 1983
A:Title: Covalent structure of chicken pepsinogen.
A:Reference number: A00984; MUID:84004412; PMID:6617663
A:Accession: A00984
A:Molecule type: protein
A:Residues: 16-87, 'S', 89-382 <BAU>
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; gastric juice; glycoprotein; hydrolase; protein digestion
F:16-57/Domain: activation peptide #status experimental <AP>
F:58-382/Product: pepsin A #status predicted <MAT>
F:92,275/Active site: Asp #status predicted
F:105-110,266-270,305-338/Disulfide bonds: #status experimental
F:128/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 12.4%; Score 301; DB 1; Length 382;
Best Local Similarity 24.0%; Pred. No. 5e-17;
Matches 88; Conservative 68; Mismatches 126; Indels 84; Gaps 13;
QY 30 YVEMTVGSPQTLNLIQVDTGSSNFAVCAAPHPFL-----HRYVQRLSSYVRLRKG 82
DB 74 YGTTISIGTQQDFTVDFGSSMLV---PSIYCKSSACSNHKKRDPKSSYVSTNET 130
QY 83 VYVYTGQKWEGLGTDLVSIHPGNVTVVRANIAITESDK-FFINGSNWEGILGLAYAE 141
DB 131 VYIAYGTGSMGILGYDTAV---SSIDVQVQIFGLSETEPGSFYFCNFDGILGLAFPS 187
QY 142 IARPDLSLEFPFSLVQKTHV-PNLFSQLGCGAGPLNSEVLASVGGSMIIGIDHSLY 200
DB 188 IS--SSGATPVDFNMNSQHLVAQDLFSYLSKDG-----ETGSFVLFGGIDPNYT 235
QY 201 TGSWLTPIRREYVEVLIIVRVEINGDOLK--MDCKEYNVDKSIDVSGTTLNRLPKKVEE 258
DB 236 TKGIYVWPLSAEYVQWQTMDRVTGNKYVACFFTC-----QAIVDTGTSLLVMPQAYN 289
QY 259 AAVKSIKAASSTE-----KFPDGFNLGEQLVQWQAGTPWNIPFVSIYLMGEVTNQS 311
DB 290 RIKDLGVSSDGEISCDISKLPD-----VTFHINGHA--- 322
QY 312 FRITILPQOYLRVEDVATSDQCKYFAISOSTGT-----VMGAVINEGFVYVDFRAR 365
DB 323 -----FTLPASAYVLNEDGSCMLGFENMGTPTELGEQWILGDVFIREYVYIFDRAN 373
QY 366 KRIGFA 371
DB 374 NKVGLS 379
RESULT 8
A34401
cathepsin E (EC 3.4.23.34) precursor - human
C:Species: Homo sapiens (man)
C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 22-Jun-1999
C:Accession: A42038; A34401; S35663; S34467; A34643; B34643
R:Azuma, T.; Liu, W.; Vander Laan, D.J.; Bowcock, A.M.; Taggart, R.T.
J. Biol. Chem. 267, 1609-1614, 1992
A:Title: Human gastric cathepsin E gene. Multiple transcripts result from alternative pro
A:Reference number: A42038; MUID:92112877; PMID:1370478
A:Accession: A42038

A:Molecule type: DNA
A:Residues: 1-396 <AZU>
A:Cross-references: GB:M82847; NID:q181203; PIDN:AAA52300.1; PID:g18120
A:Note: sequence extracted from NCBI backbone (NCBIN:75963, NCBIN:75966, NCBIN:759
J. Azuma, T.; Pals, G.; Mohandas, T.K.; Couvreur, J.M.; Taggart, R.T.
J. Biol. Chem. 264, 16748-16753, 1989
A:Title: Human gastric cathepsin E. Predicted sequence localization to chromosome
A:Reference number: A34401; MUID:89380302; PMID:2674141
A:Accession: A34401
A:Molecule type: mRNA
A:Residues: 1-396 <AZ2>
A:Cross-references: GB:J05036; NID:q181193; PIDN:AAA52130.1; PID:g181194
R:Takeda-Ezaki, M.; Yamamoto, K.
Arch. Biochem. Biophys. 304, 352-358, 1993
A:Title: Isolation and biochemical characterization of procathepsin E from human e
A:Reference number: S35663; MUID:93349047; PMID:8346912
A:Accession: S35663
A:Status: preliminary
A:Molecule type: protein
A:Residues: 20-38; 54-76 <TAK>
R: Hill, J.; Montgomery, D.S.; Kay, J.
FEBS Lett. 326, 101-104, 1993
A:Title: Human cathepsin E produced in E. coli.
A:Reference number: S34467; MUID:93314762; PMID:8325357
A:Accession: S34467
A:Status: preliminary
A:Molecule type: protein
A:Residues: 57-60, 62-81 <HIL>
R: Athauda, S.B.P.; Matsuzaki, O.; Kageyama, T.; Takahashi, K.
Biochem. Biophys. Res. Commun. 168, 878-885, 1990
A:Title: Structural evidence for two isozymic forms and the carbohydrate attachmen
A:Reference number: A34643; MUID:90241267; PMID:2334440
A:Accession: A34643
A:Status: preliminary
A:Molecule type: protein
A:Residues: 54-58, 'XX', 62-64, 'M', 66-89, 'X', 91-95 <ATH>
A:Accession: B34643
A:Status: preliminary
A:Molecule type: protein
A:Residues: 54-59, 'X', 61-68 <AT2>
C:Genetics:
A:Gene: GDB:CTSE
A:Cross-references: GDB:119821; OMIM:116890
A:Map position: 1q31-1q31
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; blocked amino end; hydrolase; zymogen
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-53/Domain: activation peptide #status predicted <PRO>
F:54-396/Product: cathepsin E #status predicted <MAT>
F:18/Modified site: blocked amino end (Gln) (in mature form) (probably pyrrolidone
F:96,281/Active site: Asp #status predicted
Query Match 12.4%; Score 299.5; DB 2; Length 396;
Best Local Similarity 25.9%; Pred. No. 7e-17;
Matches 100; Conservative 65; Mismatches 148; Indels 73; Gaps 16;
QY 3 DEEPEPCRGSGFVEMVDNLKRGKSGQGVVEMTPPOTLNLIVDTGSSNFAVCA---- 58
DB 63 DQSAKEP-----LNYLD-----MEVFGTISIGSPQNFVTFDTGSSNLWVPSVYCT 110
QY 59 APHPFLHRYVQRLSSYVTVRDLRKGVVYVPTQGWEGELGTDLVSIHPGNVTVVRANIAA 118
DB 111 SPACKTHSRFPQSSTSYSQPCQSFSIQVGTGSLSLIGADQVSV-BGLTVVGQDFGKSV 169
QY 119 TESDKFFINGSNWEGILGLAYAEIARPDLSLEPFFDLSLVKQTHVNLFSQLCGAGFFLN 178
DB 170 TEPGQTFVD-AEFGILGLGLYPSLA--VGGVTPVDFDNMAQ----NLVDLPMFVSVMSSN 222
QY 179 QSEVLASVGGSMIIGIDHSLYTGSWLTPIRREYVEVLIIVRVEINGDOLKMPCKEYNY 238
DB 223 PE---CGAGSELIFGGYDHSFSGSLNWNVPVTKQAYWQIALDNIQVGG--TVMFCSE--G 275
QY 239 DKSIVDSGTTNLRLPKPKVFEAAVKSIAASSTSEKFPDGFNLGEQLVQWQAGTTPWNIPV 298

QY 351 AVIMEGFYVDFDRARKRIGFA 371
| : : | | | | : : |
Db 365 DVFIRQYFTVFDRANNQVGLA 385

Search completed: March 4, 2003, 10:01:28
Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: March 4, 2003, 10:00:56 ; Search time 14 Seconds
(without alignments)
1350.943 Million cell updates/sec

Title: US-09-723-722A-43

Perfect score: 2419

Sequence: 1 ETDEPEPRGRSGFVEMVD.....CLRLRQHQHDFADDSILK 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2419	100.0	501	1	BACE_HUMAN
2	2392	98.9	501	1	BACE_RAT
3	2390	98.8	501	1	BACE_MOUSE
4	1171.5	48.4	518	1	BACE2_HUMAN
5	327	13.5	324	1	PEP1_GADMO
6	314.5	13.0	390	1	CATD_BOVIN
7	306	12.6	387	1	PEP1_RABIT
8	302.5	12.5	388	1	PEP4_MACFU
9	302	12.5	367	1	PEPA_CHICK
10	299.5	12.4	396	1	CATE_HUMAN
11	298.5	12.3	383	1	PEPE_CHICK
12	296	12.2	387	1	PEP2_RABIT
13	296	12.2	391	1	CATE_CAVPO
14	295.5	12.2	412	1	CATD_HUMAN
15	295	12.2	387	1	PEP4_RABIT
16	289.5	12.0	388	1	PEP2_MACFU
17	287	11.9	407	1	CATD_RAT
18	286	11.8	387	1	PEP3_RABIT
19	285.5	11.8	398	1	CATE_RAT
20	285	11.7	398	1	CATE_CHICK
21	283.5	11.7	388	1	PEPA_MACMU
22	282.5	11.7	388	1	PEPA_HUMAN
23	281.5	11.6	381	1	CHYM_SHEEP
24	281.5	11.6	388	1	PEP1_MACFU
25	279.5	11.6	410	1	CATD_MOUSE
26	278.5	11.5	386	1	PEPA_PIG
27	278	11.5	387	1	PEPA_CALJA
28	277.5	11.5	396	1	CATE_RABIT
29	277.5	11.5	397	1	CATE_MOUSE
30	276.5	11.4	419	1	CARV_CANAL
31	273.5	11.3	381	1	CHYM_BOVIN
32	273	11.3	388	1	PEPF_RABIT
33	270.5	11.2	377	1	PEPC_MACFU

RESULT 1

ID	BACE_HUMAN	STANDARD:	PRT:	501 AA.
AC	P56817: Q9JUT5; Q9BYC1: Q9BYB9;			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)			
DE	(Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl			
DE	protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)			
DE	(Memapsin-2).			
GN	BACE OR BACE1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RC	TISSUE=Brain;			
RX	MEDLINE=20002972; PubMed=10531052;			
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,			
RA	Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,			
RA	Fisher S., Fuller J., Edenson S., Lille J., Jarosinski M.A.,			
RA	Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,			
RA	Treanor J., Rogers G., Citron M.;			
RT	"Beta-secretase cleavage of Alzheimer's amyloid precursor protein by			
RL	the transmembrane aspartic protease BACE.";			
RL	Science 286:735-741(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND			
RP	CHARACTERIZATION.			
RC	TISSUE=Brain;			
RX	MEDLINE=20051717; PubMed=10591214;			
RA	Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavello R.,			
RA	Davis D., Doan M., Dorey H.F., Frigon N., Hong J., Jacobson-Croak K.,			
RA	Jewett N., Keim P., Knops J., Lieberburg I., Power W., Tan H.,			
RA	Tatsuno G., Tung J., Schenk D., Seubert P., Suomensari S.M., Wang S.,			
RA	Walker D., Zhao J., McConlogue L., Varghese J.;			
RT	"Purification and cloning of amyloid precursor protein beta-secretase			
RL	from human brain.";			
RL	Nature 402:537-540(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RX	MEDLINE=20051717; PubMed=10591213;			
RA	Van R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,			
RA	Brasier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,			
RA	Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;			
RT	"Membrane-anchored aspartyl protease with Alzheimer's disease beta-			
RL	secretase activity.";			
RL	Nature 402:533-537(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RX	MEDLINE=20120043; PubMed=10656250;			
RA	Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,			
RA	Chapman C., Glover I.S., Murphy K.E., Southan C.D., Ryan D.M.,			
RA	Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;			

Q28057 bos taurus
Q9n2d2 callithrix
Q9m2s8 ovis aries
P07267 saccharomyc
Q01294 neurospora
Q9n2d3 callithrix
Q64411 cavia porce
Q03168 aedes aegypt
P20142 homo sapien
Q28389 equus cabal
P00795 sus scrofa
P42211 oryza sativ

ALIGNMENTS

1

"Identification of a novel aspartic proteinase (Asp 2) as beta-secretase.";
Mol. Cell. Neurosci. 14:419-427(1999).
[5]
SEQUENCE FROM N.A. (ISOFORM B).
TISSUE=Brain, and pancreas;
RA Michael B., De Pietri Tonelli D., Zaccchetti D., Keller P.;
RT "New beta-site APP cleaving enzyme isoform (BACE-1b) obtained from
human brain and pancreas.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE FROM N.A. (ISOFORM C).
TISSUE=Pancreas;
RC Zaccchetti D., De Pietri Tonelli D., Schnurbus R.;
RA "New beta-site APP cleaving enzyme isoform (BACE-1c) obtained from
human pancreas.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE FROM N.A. (ISOFORMS B; C AND D).
TISSUE=Brain;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site
amyloid precursor protein cleaving enzyme (BACE) and their effect on
amyloid beta-peptide production.";
RL Neurosci. Lett. 307:9-12(2001).
[8]
SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
MEDLINE=20144050; PubMed=10677483;
RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
beta-amyloid precursor protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
[9]
DISULFIDE BONDS.
MEDLINE=21950860; PubMed=11953458;
RA Fischer F., Molinari M., Bodendorf U., Paganetti P.;
RT "The disulfide bonds in the catalytic domain of BACE are critical but
not essential for amyloid precursor protein processing activity.";
RL J. Neurochem. 80:1079-1088(2002).
CC -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 4 isoforms; A/BACE-1A/BACE-501 (shown here),
B/BACE-1B/BACE-I-476; C/BACE-1C/BACE-I-457 and D/BACE-1D/BACE-I-
452; are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL AF190725; AAF04142.1; -
DR EMBL AF201468; AAF18982.1; -
DR EMBL AF200343; AAF17079.1; -
DR EMBL AF204943; AAF26367.1; -
DR EMBL AF338816; AAK38374.1; -
DR EMBL AF338817; AAK38375.1; -
DR EMBL AB050436; BAB40931.1; -
DR EMBL AB050437; BAB40932.1; -
DR EMBL AB050438; BAB40933.1; -
DR EMBL AF200193; AAF13715.1; -
DR HSSP: P32329; LYPS. -
DR MEROPS; A01.004; -.

DR Genew: HGNC:933; BACE.
DR MIM: 604252; -.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW Signal; Alternative splicing.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 45
FT CHAIN 46 501 BETA-SECRETASE.
FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 458 478 POTENTIAL.
FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 289 289 BY SIMILARITY.
FT DISULFID 216 420
FT DISULFID 278 443
FT DISULFID 330 380
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 146 189 MISSING (IN ISOFORM C AND ISOFORM D).
FT VARSPPLIC 190 214 MISSING (IN ISOFORM B AND ISOFORM D).
SQ SEQUENCE 501 AA; 55763 MW; 377CE4C824ACEF05 CRC64;

Query Match 100.0%; Score 2419; DB 1; Length 501;
Best Local Similarity 100.0%; Pred. No. 2.5e-193;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEEPCGRGSGFVEMVDNLGRKGSGGYVYVEMTVGSPPTNLINLVDTGSSNFAVGAAP 60
DB 46 ETDEEPEEPCGRGSGFVEMVDNLGRKGSGGYVYVEMTVGSPPTNLINLVDTGSSNFAVGAAP 105

QY 61 HPFLHRYQRLSSTYRDLRGVYVYPTGKWEGLGTDLVSIHPGPNVTVRANIAITE 120
DB 106 HPFLHRYQRLSSTYRDLRGVYVYPTGKWEGLGTDLVSIHPGPNVTVRANIAITE 165

QY 121 SDRFFINGSNWEGTILGLAYAEIARPDSDLPEPFDSLVKOTHVPLNLSLQLCGAGFPLNOS 180
DB 166 SDRFFINGSNWEGTILGLAYAEIARPDSDLPEPFDSLVKOTHVPLNLSLQLCGAGFPLNOS 225

QY 181 EVLASVGSMTIGGIDHSYLTGSLWYTPTRREWYVEVIVRVEINGDLKMDCKEYNYDK 240
DB 226 EVLASVGSMTIGGIDHSYLTGSLWYTPTRREWYVEVIVRVEINGDLKMDCKEYNYDK 285

QY 241 SIYDSGTNLRPKKVFEEAAVKSIAASSTKFPDGFNLGEQLVQVQAGTTPWNIFPVIS 300
DB 286 SIYDSGTNLRPKKVFEEAAVKSIAASSTKFPDGFNLGEQLVQVQAGTTPWNIFPVIS 345

QY 301 LYLMGEVTNOSFRITILPQOYLRLPVEDVATSDDDCYKFAISQSSTGTVMGAVIMEGFVYV 360
DB 346 LYLMGEVTNOSFRITILPQOYLRLPVEDVATSDDDCYKFAISQSSTGTVMGAVIMEGFVYV 405

QY 361 FDRARRKIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPTQDESLMTIAYVMAAI 420
DB 406 FDRARRKIGFAVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPTQDESLMTIAYVMAAI 465

QY 421 CALFMPLPLCLMVCMWRCLRLRQOHHDDFADDISLLK 456
DB 466 CALFMPLPLCLMVCMWRCLRLRQOHHDDFADDISLLK 501

RESULT 2
BACE_RAT
ID BACE_RAT STANDARD: PRT: 501 AA.
AC P56819;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-secretase precursor (EC 3.4.23.-) (beta-site APP cleaving enzyme)

FT	CARBOHYD	241	241	N-LINKED (GLCNAC...)	(POTENTIAL).
SQ	SEQUENCE	390 AA:	42488 MW:	5B38AALC3C48D35	CRC64:
	Query Match		13.0%:	Score 314.5:	DB 1: Length 390:
	Best Local Similarity		28.0%:	Pred. No. 1.2e-18:	
	Matches 113:	Conservative 72:	Mismatches 128:	Indels 91:	Gaps 21:
QY	8	EPG-RRGSFVEMVDNLRGSGGQGYVYEMTVGSPPTLNLVDTGSSNFAVGAAPHPPL--	64		
DB	39	EPAVRQGPTELLKKNYMDAQ---YGEIGITGTPQCFTVFVDTGSAANLWVPSIHCKLLDI	95		
QY	65	---HRYQRQLSSYYRDLRKGYV--VPYTGQKWEGLGTDLVSLPHGPN-----VTVR	112		
DB	96	ACWTHRKYNDSKSTY-VKNGTFTDIHYGSGSLGYSQDTSVSPCNPSSSSPGGVTQ	153		
QY	113	ANI--AATESDKFFGNSWEGILGAYAEIARPDSDLEPPFDSLVKQTHV--PNLFSLO	169		
DB	154	ROTFGEALIQGVVFI-AAKFDGILGWAYPRIS--VNNLPVFDNLMOKLVKDKNVS--	208		
QY	170	LCGAGFPLNQSEVLASVGSGSMITGGIDHSLYTGSLWYTPTRREWYVEYIIVRVEINGDL	229		
DB	209	----FFLNR-DPKAQPGGELMLGGTDSKYRGSLMFPHNVTRQAYWQIIMDQLDV-CSSL	261		
QY	230	KMDCKEYNYDKSIVDSGTTNRLRPLPKKVFEEAAVKSIKKAASSTKFPDGFGLGEOLV-CWQA	288		
DB	262	TV-CK--GCCEAIVDTGTSUIVGPVEEVRELQKAIGAVPLIQ-----GEYMIPEKV	310		
QY	289	GTTTPNFIPIVSLYLMEVNTNQSFRITILPQOYLRPVEDVATSDODCYKFAISOSSGTG-	347		
DB	311	SS-----LPEVTVKLG-----KDYALSPED-VALKVSOAETTV	344		
QY	348	-----VMGAVIMEGFYVYVDFRARKRIGFAVSA	374		
DB	345	LSGFMGMDIPPPGGPLWILGDVPIGRYTYTFVDRDQNRVGLAEAA	388		
RESULT 7	PEPI_RABIT	STANDARD:	PRT:	387 AA.	
ID	PEPI_RABIT				
AC	P28712:				
DT	01-DEC-1992 (Rel. 24, Created)				
DT	01-DEC-1992 (Rel. 24, Last sequence update)				
DE	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Pepsin II-1 precursor (EC 3.4.23.1) (Pepsin A).				
OS	Oryctolagus cuniculus (Rabbit).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.				
ON	NCBI_Taxid=9986;				
RX	SEQUENCE FROM N.A.				
RP	MEDLINE=91009127; PubMed=2129536;				
RA	Kageyama T., Tanabe K., Koiwai O.;				
RT	"Structure and development of rabbit pepsinogens. Stage-specific				
RT	zymogens, nucleotide sequences of cDNAs, molecular evolution, and				
RT	gene expression during development."				
RL	J. Biol. Chem. 265:17031-17038(1990).				
CC	-I- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS				
CC	INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE				
CC	ALSO CLEAVED TO SOME EXTENT.				
CC	-I- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably				
CC	aromatic, residues in P1 and P1' positions. Cleaves 1-Phe-1-Val-2,				
CC	4-Gln-1-His-5, 13-Glu-1-Ala-14, 14-Ala-1-Leu-15, 15-Leu-1-Tyr-16,				
CC	16-Tyr-1-Leu-17, 23-Gly-1-Phe-24, 24-Phe-1-Phe-25 and 25-Phe-1-				
CC	Tyr-26 bonds in the B chain of insulin.				
CC	-I- DEVELOPMENTAL STAGE: PEPsinogens IN GROUP I, II, and III WHERE				
CC	THE PREDOMINANT ZYMOGENS AT LATE POSTNATAL STAGE.				
CC	-I- MISCELLANEOUS: THE EXPRESSION OF PEPsinogen GENES IS REGULATED BY				
CC	HORMONES AND RELATED SUBSTANCES.				
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.				
DR	PIR: B38302; B38302.				
DR	HSSP: P00791; IPSA.				
DR	MEROPS: A01.001; -.				
DR	InterPro: IPR001461; AsproteaseA1				

```

DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; Zymogen; Signal;
KW Phosphorylation; Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 59
FT CHAIN 60 387
FT MOD_RES 129 129
FT ACT_SITE 93 93
FT ACT_SITE 276 276
FT DISULFID 106 111
FT DISULFID 267 271
FT DISULFID 310 343
SQ SEQUENCE 387 AA; 42070 MW; A6EC48F715541A48 CRC64;

Query Match 12.6%; Score 306; DB 1; Length 387;
Best Local Similarity 27.1%; Pred. No. 6.3e-18;
Matches 98; Conservative 67; Mismatches 131; Indels 66; Gaps 15;

QY 30 YVEMTVGSPPTNLNILDVTCSSNFVAG---AAPHPLHRYQORQLSSYRDLRKGVYV 85
DB 75 YFGTISTGTPPEPTVIFDTGSSNLWPFSTYCSLACFLHKKRPNDDSTFQATSETLSI 134
QY 86 PYTGKWEGLGTDLSIPHGPNNVTVRANIAAITESED---KFFINGSNWEGILGLAYAEI 142
DB 135 TYGTGSMTGILGYTVAV---GNIEDTNOIFGLSKTEPGITFLV--APFDGILGLAYPSI 189
QY 143 ARPDSDLEPFDSLVKQTHV--PNLFSIOLCGAGPLNQSEVLASVGGSMIIGDHSLYT 201
DB 190 SASDAT--PVFDNMWNEGLVSEDLFSVYLSNG-----BKGSVMVFGGIDSSVYT 237
QY 202 GSLWYTPIREWYEVLIIVRVEINGQDLKM--DCKEYNDKSIYDSGTTNLRPKKVFEEA 259
DB 238 GSNLWVPVSHGYYQWITMDSITINGETIACADSC-----QAVVDGTGSLLAGPTSAISK 291
QY 260 AVKSAKASSTKEKPPDGFVWGLQV--CWQACTTPWNIPFVLSYLMGEVTVNQSFRIILP 318
DB 292 IQSYIGASKNL-----LGENIISCAIDSLPDIVF-----FINN 325
QY 319 QQYLRPVED-VATSDQDC---YKFAISOSSTGT---VMGAVIMEGFYVVDRAKRIGFAV 372
DB 326 VQYPLPASAVILKDDCLSGFGMNLDTSYGELWILGDVEIRQYFTVFDRAANNQVGLAA 385
QY 373 SA 374.
DB 386 AA 387

RESULT 8
PEP4_MACFU STANDARD; PRT; 388 AA.
AC P27678;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pepsin A-4 precursor (EC 3.4.23.1) (Pepsin I/II).
GN PGA.
OS Macaca fuscata fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9543;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-70.
RC TISSUE=Gastric mucosa;
RC MEDLINE=92037645; PubMed=1935977;
RA Kageyama T., Tanabe K., Koizumi O.;
RT "Development-dependent expression of isozymogens of monkey
RL pepsinogens and structural differences between them.";
RL Eur. J. Biochem. 202:205-215(1991).
CC -!- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS

```

```

CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably
CC aromatic, residues in P1 and P1' positions. Cleaves L-Phe-L-Val-2,
CC 4-Gln-L-His-5, 13-Glu-L-Ala-14, 14-Ala-L-Leu-15, 15-Leu-L-Tyr-16,
CC 16-Tyr-L-Leu-17, 23-Gly-L-Phe-24, 24-Phe-L-Phe-25 and 25-Phe-L-
CC Tyr-26 bonds in the B chain of insulin.
CC -!- MISCELLANEOUS: THE EXPRESSION OF PEPsin GENES IS REGULATED BY
CC HORMONES AND RELATED SUBSTANCES.
CC -!- MISCELLANEOUS: EACH PEPsin IS CONVERTED TO CORRESPONDING
CC PEPsin AT PH 2.0 IN PART AS A RESULT OF THE RELEASE OF A 47 AA
CC ACTIVATION SEGMENT AND IN PART AS A RESULT OF STEPWISE PROTEOLYTIC
CC CLEAVAGE VIA AN INTERMEDIATE FORM(S).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X59753; CAA42425.1;
CC PIR; S16065; S16065.
CC PIR; S19682; S19682.
CC HSP; P00790; IFSN.
CC MEROPS; A01.001;
CC InterPro; IPR001461; AspproteaseA1.
CC InterPro; IPR001969; Aspprotease_site.
CC InterPro; IPR001791; Laminin_G.
CC Pfam; PF00026; asp. 1.
CC PRINTS; PR00792; PEPsin.
CC PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; Zymogen; Multigene family;
KW Signal; Glycoprotein.
FT SIGNAL 1 15
FT PROPEP 16 59
FT PROPEP 39 62
FT CHAIN 63 388
FT ACT_SITE 94 94
FT ACT_SITE 277 277
FT DISULFID 107 112
FT DISULFID 268 272
FT DISULFID 311 344
FT CARBOHYD 88
SQ SEQUENCE 388 AA; 41955 MW; A2923AB1F7FCDEB9 CRC64;

Query Match 12.5%; Score 302.5; DB 1; Length 388;
Best Local Similarity 27.6%; Pred. No. 1.2e-17;
Matches 107; Conservative 64; Mismatches 133; Indels 83; Gaps 17;

QY 3 DEEPPEPGRGSRFVEMVDNLKRGSGQGYVEMTVGSPPTNLNILDVTCSSNFVAGAAHP 62
DB 64 DEQPLE-----NYLDV-----EYFGTIGTGAQNETVVDFTGSSNLW---PSV 105
QY 63 FL-----HRYQORQLSSTYRDLRKGVYVPTGCKWSEGLGTDLSVPHGPNVTVRANI 115
DB 106 YCYSLACMDHNLFPQDSSTYRATSKVTSITYGTGSMTGILGYDFVKV---GGISDTNQI 162
QY 116 AAITESDK-PFINGSNWEGLGLAYAEIARPDSDLEPFDSLVKQTHV--PNLFSIOLCGA 173
DB 163 FGLSETEPGFLYFAPFDGILGLAYPSIS--SSGATPVFDNINWQRLVSQDLFSVYLSAD 220
QY 174 GFPLNQSEVLASVGGSMIIGDHSLYTGLSLWYTPIREWYEVLIIVRVEINGQDL--KM 231
DB 221 ----DQS-----GSVIFGGIDSSYTGSLNWWVPSVVEGYWQISVDSITMNGKTIACAK 270
QY 232 DCKEYNDKSIYDSGTTNLRPKKVFEEAASIKASSTKEKPPDGFVWGLQV--CWQACT 290
DB 271 GC-----QAVVDGTGSLTGTPTSIANIQSDIGASENSD-----GEMVVSATISS 316
QY 291 TPWNIFPVISLYLMGEVTVNQSFRIILPQQY--LRPVEDVATSDQDCYK-----FAISQSS 344

```

```
Db 317 LPDIVF-----TINGVOYPLPPSAVILQSGSCTSGFOGMDVPTESG 358
QY 345 TGTVMGAVMEGFYVVFDRKRIGFA 371
Db 359 ELWILGDVFIROYFTVFDNRANNOVGLA 385

RESULT 9
PEPA_CHICK
ID PEPA_CHICK STANDARD: PRT: 367 AA.
AC P00793;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pepsin A precursor (EC 3.4.23.1).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=84004412; PubMed=6617663;
RA Baudys M., Kostka V.;
RT "Covalent structure of chicken pepsinogen.";
RL Eur. J. Biochem. 136:89-99(1983).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY. ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably
CC aromatic, residues in p1 and p1' positions. Cleaves 1-Phe-[Val]-2,
CC 4-Gln-[His]-5, 13-Glu-[Ala]-14, 14-Ala-[Leu]-15, 15-Leu-[Tyr]-16,
CC 16-Tyr-[Leu]-17, 23-Gly-[Phe]-24, 24-Phe-[Phe]-25 and 25-Phe-[Tyr]-26 bonds in the B chain of insulin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR PIR; A00984; PECH.
DR HSSP; P00794; 4CMS.
DR MEROPS; A01.0P.1; -.
DR InterPro; IPR001461; AsparticaseA1.
DR InterPro; IPR001969; Asparticase_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; Zymogen; Glycoprotein.
FT PROPEP 1 42 ACTIVATION PEPTIDE.
FT CHAIN 43 367 PEPsin A.
FT ACT_SITE 77 77
FT ACT_SITE 260 260
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .).
FT DISULFID 90 95
FT DISULFID 251 255
FT DISULFID 290 323
SQ SEQUENCE 367 AA; 40431 MW; 0C547E7FD8F5B341 CRC64;

Query Match 12.5%; Score 302; DB 1; Length 367;
Best Local Similarity 24.0%; Pred. No. 1.3e-17;
Matches 88; Conservative 69; Mismatches 125; Indels 84; Gaps 13;

QY 30 YVYVMTGSPQTLNVLVDGSSNFAVGAHPFL-----HRYQRLSSSTYRDLRKG 82
Db 59 YGPISTGTPOQDFSVIFDTGSSNLWV---PSYICKSSACSNHKKRFPDPSKSTYVSTNET 115
QY 83 YVYVPTGKGEGELGTLVLSIPHPNVTVRANIAATESDK -FFINGSNNEGILGLAYAE 141
Db 116 VYIAYGTGSMGILGYDTAVV---SSIDVQNIQIFGUSETEPGGFFYYCNDGILGLAPPS 172
QY 142 IARPDSDLPEPFDLSLVKQTHV--PNLFSLOLCGAGFLPNQSEVLASVGGSMIIGIDHSLY 200
Db 173 IS--SSGATPVFDNMMSQHLVAQDLFSVYLSKDG-----ETSGFVLFGGIDPNYT 220
QY 201 TGSWYTPITREYWEYVIVIRVEINGQDLK--MDCKEYNDKSIIVSDGNTNLRPLPKKVF 258
```

```
Db 221 TKGIYVWPLSAETYYQWITMDRVTVGKYNKYVACFFTC-----QAIVDTGTSLLVMPQGAYN 274
QY 259 AAVKSIIKAASTPE-----KRPDGEWLGBQLVCWQACTTPWNIFPVLISLYLMGEVTNQS 311
Db 275 RIIRKDLGVSSDGEISCDISKLDP-----VTFHINGHA----- 307
QY 312 FRITILPOQYLRPVEDVATSQDDCYKFAISQSSGTG-----VMGAVIMEGFYVVFDRAR 365
Db 308 -----FTLPASAYVLNEDGSCMLGFENMGTPTELGEOWILGDVPIREYVYVIFDRAN 358
QY 366 KRIGFA 371
Db 359 NKVGLS 364

RESULT 10
CATE_HUMAN
ID CATE_HUMAN STANDARD: PRT: 396 AA.
AC P14091;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin E precursor (EC 3.4.23.34).
GN CTSE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89380302; PubMed=2674141;
RA Azuma T., Pals G., Mohandas T.K., Couvreur J.M., Taggart R.T.;
RT "Human gastric cathepsin E. Predicted sequence, localization to
RT chromosome 1, and sequence homology with other aspartic
RT proteinases.";
RL J. Biol. Chem. 264:16748-16753(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112877; PubMed=1370478;
RA Azuma T., Liu W.G., Vander Laan D.J., Bowcock A.M., Taggart R.T.;
RT "Human gastric cathepsin E gene. Multiple transcripts result from
RT alternative polyadenylation of the primary transcripts of a single
RT gene locus at 1q31-q32.";
RL J. Biol. Chem. 267:1609-1614(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Tatnell P.J., Kay J.;
RT "Human procathepsin E.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 54-68; 77-95; 141-155; 275-285 AND 389-396.
RX MEDLINE=90241267; PubMed=2334440;
RA Ahnada S.B.P., Matsuzaki O., Kgeyama T., Takahashi K.;
RT "Structural evidence for two isozyme forms and the carbohydate
RT attachment site of human gastric cathepsin E.";
RL Biochem. Biophys. Res. Commun. 168:878-885(1990).
CC -1- FUNCTION: DUE OF ITS INTRACELLULAR LOCATION AND DISTRIBUTION IN
CC LYMPHOID ASSOCIATED TISSUE, IT MAY HAVE A ROLE IN IMMUNE FUNCTION.
CC -1- CATALYTIC ACTIVITY: Similar to cathepsin D, but slightly broader
CC specificity.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M84424; AAA52300.1; -.
CC EMBL; M84413; AAA52300.1; JOINED.
DR
```

```
DR EMBL; M84417; AAA52300.1; JOINED.
DR EMBL; M84418; AAA52300.1; JOINED.
DR EMBL; M84419; AAA52300.1; JOINED.
DR EMBL; M84420; AAA52300.1; JOINED.
DR EMBL; M84421; AAA52300.1; JOINED.
DR EMBL; M84422; AAA52300.1; JOINED.
DR EMBL; J05036; AAA52130.1; -.
DR EMBL; AJ250717; CAB82850.1; -.
DR PIR; A34401; A34401.
DR PIR; A34643; A34643.
DR PIR; A42038; A42038.
DR HSP; P00794; 4CMS.
DR MEROPS; A01.010; -.
DR Genew; HGNC:2530; CTSE.
DR MIM; 116890; -.
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 17
FT PROPEP 18 53 ACTIVATION PEPTIDE.
FT CHAIN 54 396 CATHEPSIN E.
FT MOD_RES 18 18 PYRROLIDONE CARBOXYLIC ACID.
FT ACT_SITE 96 96 BY SIMILARITY.
FT ACT_SITE 281 281 BY SIMILARITY.
FT DISULFID 60 60 INTERCHAIN (PROBABLE).
FT DISULFID 109 114 BY SIMILARITY.
FT DISULFID 272 276 BY SIMILARITY.
FT DISULFID 314 351 BY SIMILARITY.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .).
FT CARBOHYD 220 220 O-LINKED (POTENTIAL).
FT CARBOHYD 333 333 O-LINKED (POTENTIAL).
SQ SEQUENCE 396 AA; 4793 MW; 40B643C5FB01521E CRC64;

Query Match 12.4%; Score 299.5; DB 1; Length 396;
Best Local Similarity 25.9%; Pred. No. 2.2e-17;
Matches 100; Conservative 65; Mismatches 148; Indels 73; Gaps 16;

QY 3 DEEPEPGRGSGFVEMVDNLGRSGGQYVEMTVGSPQTLNLLVDTGSSNFAVGA---- 58
D 63 DOSAKEP-----LNYLD-----MEYFGTISGSPQNTVFDTGSSNFWPSVYCT 110
QY 59 APHPLHRYQRLSSRYRDLKGVVYPTQGWKEGELGTLVSIHPGPNVTVRANIAAI 118
D 111 SPACKTHSRFQPSQSSSYQPGQSFQYGTGSLGIIGADQVSV-EGLTVVGQGFGEV 169
QY 119 TESDKFFINGSNMGILGLAYAEIARPDSDSLEPPFDSLVKQTHVPNLFSLQCGAGPPLN 178
D 170 TEPQTEVD-AEFDGILGLGVPSLA--VGGVTPVFDNMMAQ---NLVDLPMSVYKSN 222
QY 179 QSEVLASVGSMIIGGDHSLYTGSLWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNY 238
D 223 PE---GGAGSELIFGGYDHGFSGLSNWVPTKQYQALDNIQVGG--TVMFCSE--G 275
QY 239 DKSIVDSGTTNLRPLKPKVFAAANKVSKAASSTERKFPDGEMLGELQVLCVQAGTTPWTFPV 298
D 276 QCAIVDTGTSLLITGSPDKIKQLQNAICAP-----VGEYAVE-----CANLNWPD 322
QY 299 ISLYLMGEVNTQSPRITILPQQYLRPVEDVATSDODCKYKFAISQSSSTG----- 346
D 323 VTFTING-----VPYTLSPATY--TLLDVFDGMQFC-----SSGFGQLDHPAPG 366
QY 347 -TVMGVIMGEFYVFDPRARKRIGFA 371
D 367 LWILGDVFIQFYSVFDRGNRRVGLA 392

RESULT 11
PEPE_CHICK
ID PEPE_CHICK STANDARD; PRT; 383 AA.
AC P16476;
```



```
Db 295 AVGANQNT-----YGEYSV-----NCSHILAMPDVVFVIGGI----- 326
QY 319 QOYLRPVEDVA---TSDDCYKFAISQSTGTGVMGAVIMEGEVVFVDRARKRIGFA 371
Db 327 -QY--PVPALAYTEQNGOGCTMSSFSQSSADLWILGDVFRVYISFDNRANRVGLA 380

RESULT 12
PEP2_RABIT
ID PEP2_RABIT STANDARD; PRT; 387 AA.
AC F27821;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pepsin II-2/3 precursor (EC 3.4.23.1) (Pepsin A).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=91009127; PubMed=2129536;
RA Kageyama T., Tanabe K., Koiwai O.;
RT "Structure and development of rabbit pepsinogens. Stage-specific
RT zymogens, nucleotide sequences of cDNAs, molecular evolution, and
RT gene expression during development.";
RL J. Biol. Chem. 265:17031-17038(1990).
CC -!- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably
CC aromatic, residues in p1 and p1' positions. Cleaves 1-Phe-I-Val-2,
CC 4-Gln-I-His-5, 13-Glu-I-Ala-14, 14-Ala-I-Leu-15, 15-Leu-I-Tyr-16,
CC 16-Tyr-I-Leu-17, 23-Gly-I-Phe-24, 24-Phe-I-Phe-25 and 25-Phe-I-
CC Tyr-26 bonds in the B chain of insulin.
CC -!- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE
CC THE PREDOMINANT ZYMOGENS AT LATE POSTNATAL STAGE.
CC -!- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
CC HORMONES AND RELATED SUBSTANCES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M59235; AAA85369.1;
DR PIR; C38302; C38302.
DR HSP; P00790; IFSN.
DR MEROPS; A01.001;
DR InterPro; IPR001461; AsparticaseA1.
DR InterPro; IPR001969; Asparticase_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; Zymogen; Signal;
KW Phosphorylation; Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 59 ACTIVATION PEPTIDE.
FT CHAIN 60 387 PEPSIN II-2/3.
FT MOD_RES 129 129 PHOSPHORYLATION (POTENTIAL).
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 276 276 BY SIMILARITY.
FT DISULFID 106 111 BY SIMILARITY.
FT DISULFID 267 271 BY SIMILARITY.
FT DISULFID 310 343 BY SIMILARITY.
SQ SEQUENCE 387 AA; 42100 MW; 66FC331A3DC75891 CRC64;
```

Query Match

12.2%; Score 296; DB 1; Length 387;

```
Best Local Similarity 26.9%; Pred. No. 4.2e-17;
Matches 97; Conservative 63; Mismatches 135; Indels 66; Gaps 13;

QY 30 YVEMTVGSPQTLNILDVTGSSNFAVGAAPHF-----LHRYQROLSTSYRDLKKG 82
Db 75 YFGTISGTPQDFTVIFDTGSSNLW---PSTYCSLACALHRRFNPEDSSTYOGTSET 131
QY 83 YVVPYTGKGEGELGTDLVSIPIHGPNTVVRANIAATESKFFINGSNWEGTLGAYAEI 142
Db 132 LSITYGTGSMTGILGYDVKVGSIEDTNQIFGLSKTEPSLTFLF--APFDGILGLAYPSI 189
QY 143 ARPDLSLEPFDDSLVKQTHV-PNLFSLQLCGAGFPLNQSVLASVGGSMIIGGDHSLYT 201
Db 190 SSSDAT--PVFDNMNNEGLVSDQLFSVYLSSDD-----EKGLSLVMFGGIDSSVYT 237
QY 202 GSLWYTPIRREWYEVIIIVRVEINQODLKM--DCKEYNYDKSIVDSGTTNLRPKKVEA 259
Db 238 GSLNMPVSVYEGYQWQITMDSVINGETIACADSC-----QAIYDTCTSLTGP---TS 287
QY 260 AVKSIKAASSTKFPDGFGLV-CWQAGTTPWNIFPVVISLYLMGEVTVNOSFRITILP 318
Db 288 AISNIQSYIGASK-----NLGENVISCSAIDSLDIVF-----TING 325
QY 319 QOYLRPVEDVATSDDCYKFAISQSTGT-----VMGAVIMEGFYVVFDRARKRIGFAV 372
Db 326 IQYPLPASAYILKEDDDCTSGLEGMNVDTYTGELWILGDVFIQYFTVFDNRANQLGLAA 385
QY 373 S 373
Db 386 A 386

RESULT 13
CATE_CAVPO STANDARD; PRT; 391 AA.
AC P25796;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin E precursor (EC 3.4.23.34).
GN CTSE.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=92355614; PubMed=1644829;
RA Kageyama T., Ichinose M., Tsukada S., Miki K., Kurokawa K., Koiwai O.,
RA Tanji M., Yakabe E., Athauda S.B., Takahashi K.;
RT "Gastric procathepsin E and progastricins from guinea pig.
RT Purification, molecular cloning of cDNAs, and characterization of
RT enzymatic properties, with special reference to procathepsin E.";
RL J. Biol. Chem. 267:16450-16459(1992).
RN [2]
SEQUENCE FROM N.A.
RP TISSUE=Gastric mucosa;
RX MEDLINE=96073637; PubMed=8540321;
RA Kageyama T., Ichinose M., Miki K., Moriyama A., Yonezawa S.,
RA Tanji M., Athauda S.B., Takahashi K.;
RT "Isolation, characterization, and structure of procathepsin E and
RT cathepsin E from the gastric mucosa of guinea pig.";
RL Adv. Exp. Med. Biol. 362:211-221(1995).
CC -!- FUNCTION: DUE TO ITS INTRACELLULAR LOCATION AND DISTRIBUTION IN
CC LYMPHOID ASSOCIATED TISSUE, IT MAY HAVE A ROLE IN IMMUNE FUNCTION.
CC -!- CATALYTIC ACTIVITY: Similar to cathepsin D, but slightly broader
CC specificity.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


Search completed: March 4, 2003, 10:02:13
Job time : 17 secs

RL	J. Biol. Chem. 265:17031-17038(1990).	
CC	-!- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS	
CC	CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE	
CC	CC ALSO CLEAVED TO SOME EXTENT.	
CC	-!- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably	
CC	CC aromatic, residues in p1 and p1' positions. Cleaves 1-Phe- -Val-2,-	
CC	CC 4-Gln- -His-5, 13-Glu- -Ala-14, 14-Ala- -Leu-15 15-Leu- -Tyr-16,	
CC	CC 16-Tyr- -Leu-17, 23-Gly- -Phe-24, 24-Phe- -Phe-25 and 25-Phe- -	
CC	CC Tyr-26 bonds in the B chain of Insulin.	
CC	-!- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE	
CC	CC THE PREDOMINANT ZYMogens AT LATE POSTNATAL STAGE.	
CC	-!- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY	
CC	CC HORMONES AND RELATED SUBSTANCES.	
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.	
DR	PIR; D38302; D38302.	
DR	HSSP; P00790; LPSN.	
DR	MEROPS; A01.001; -.	
DR	InterPro; IPR001461; AsproteaseA1.	
DR	InterPro; IPR001969; Asprotease_site.	
DR	Pfam; PF00026; asp; 1.	
DR	PRINTS; P00792; PEPsin.	
DR	PROSITE; PS00141; ASP_PROTEASE; 2.	
KW	Hydrolase; Aspartyl protease; Digestion; Zymogen; Signal;	
KW	Phosphorylation; Multigene family.	
FT	SIGNAL 1 15	
FT	PROPEP 16 59	ACTIVATION PEPTIDE.
FT	CHAIN 60 387	PEPSIN II-4.
FT	MOD_RES 129 129	PHOSPHORYLATION (POTENTIAL).
FT	ACT_SITE 93 93	BY SIMILARITY.
FT	ACT_SITE 276 276	BY SIMILARITY.
FT	DISULFID 106 111	BY SIMILARITY.
FT	DISULFID 267 271	BY SIMILARITY.
FT	DISULFID 310 343	BY SIMILARITY.
SQ	SEQUENCE 387 AA; 42052 MW; 21ADD07782A89585 CRC64;	
	Query Match 12.2%; Score 295; DB 1; Length 387;	
	Best Local Similarity 26.1%; Pred. No. 5.le-17;	
	Matches 97; Conservative 65; Mismatches 123; Indels 86; Gaps	
Qy	30 YVEMTVGSPPTLNTLVDTGSNFAVGAAPHF-----LHRYXORQLSSTYRDLRKG 82	
	YFGTISITGTPQDFTVIFDTGSNLWV---PSTYCSSALCAHKRFNPDSSTYQGTST 131	
Db	83 VYVPYTGQKWEGLGDLVSIPHGPNVTVRANIAAITESDKFF-----INGSNWE 132	
Qy	132 LSIYGTGSGWTGLIGYDTV-----KVGSIETNOIFGLSKTEPGTLTFLPAPD 179	
Qy	133 GILGLAYAEIARPDOSLEFPDSLKVOYTHV-PNLFSLGCGGFPPLNQSEVLASVGSGM 191	

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:00:57 ; Search time 36 seconds
(without alignments)
2609.930 Million cell updates/sec

Title: US-09-723-722a-43

Perfect score: 2419

Sequence: 1 ETDEPEPGRGRSGFVEMVD.....CLRLRQHQHDFADISLKL 456

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rv1rus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2419	100.0	532	4 Q9ULS1	Q9uls1 homo sapien
2	2264.5	93.6	476	4 Q9BYC1	Q9byc1 homo sapien
3	2161	89.3	457	4 Q9BYC0	Q9byc0 homo sapien
4	2016.5	83.4	432	4 Q9BYB9	Q9byb9 homo sapien
5	1407	58.2	266	11 Q9CUU5	Q9cuu5 mus musculus
6	1160.5	48.0	439	4 Q9H2V8	Q9h2v8 homo sapien
7	1146.5	47.4	514	11 Q9JL18	Q9jl18 mus musculus
8	972.5	40.2	468	4 Q9NZL2	Q9nzl2 homo sapien
9	967.5	40.0	396	4 Q9NZL1	Q9nzl1 homo sapien
10	712.5	29.5	213	4 Q9P0D2	Q9p0d2 homo sapien
11	596.5	24.7	255	11 Q9R1P7	Q9rlp7 mus musculus
12	359.5	14.9	244	5 Q8WQY9	Q8wqy9 aphrocallis
13	332.5	13.7	391	5 Q9VKP6	Q9vxp6 drosophila
14	332	13.7	354	5 Q9GYX7	Q9gyx7 boophilus m
15	312.5	12.9	386	6 Q9BGU5	Q9bgu5 bos taurus
16	308	12.7	384	13 Q9DEC2	Q9dec2 xenopus lae

17	308	12.7	385	13 Q9DEC4	Q9dec4 rana catesb
18	305	12.6	386	6 Q9GMV7	Q9gmv7 rhinolophus
19	305	12.6	387	6 Q9GMV8	Q9gmv8 sorex ungui
20	304.5	12.6	372	5 Q9VLK3	Q9vlk3 drosophila
21	304.5	12.6	387	13 Q9DDV5	Q9ddv5 salvelinus
22	304	12.6	387	6 Q9GMV9	Q9gmv9 suncus muri
23	303.5	12.5	383	13 Q9DE45	Q9de45 salvelinus
24	303	12.5	383	13 Q9DEC3	Q9dec3 xenopus lae
25	302.5	12.5	376	13 Q9PUR8	Q9pur8 pseudopleur
26	301	12.4	382	13 Q9PRG9	Q9prg9 gallus gall
27	301	12.4	423	5 Q9VRP7	Q9vxp7 drosophila
28	298.5	12.3	384	13 Q91322	Q91322 rana catesb
29	295.5	12.2	386	6 Q9GMV6	Q9gmv6 canis fami
30	294	12.2	396	13 Q93428	Q93428 chionodraco
31	290.5	12.0	381	6 Q9GK11	Q9gk11 camelus dro
32	288	11.9	399	13 Q93458	Q93458 podarcis si
33	287.5	11.9	444	5 Q21966	Q21966 caenorhabdi
34	284	11.7	398	13 P87370	P87370 oncorhynch
35	284	11.7	427	5 P91802	P91802 schistosoma
36	281	11.6	378	13 Q9PUR9	Q9pur9 pseudopleur
37	280	11.6	390	6 Q8SQ41	Q8sq41 canis fami
38	279.5	11.6	390	6 Q9GK10	Q9gk10 camelus dro
39	278	11.5	370	6 Q9RTW1	Q9rtw1 bos taurus
40	278	11.5	399	13 Q9DD89	Q9dd89 brachydanio
41	278	11.5	422	5 Q96906	Q96906 onchocerca
42	277.5	11.5	380	6 Q28950	Q28950 sus scrofa
43	277	11.5	446	5 Q9N9H3	Q9n9h3 necator ame
44	276	11.4	389	13 Q9PWK1	Q9pwk1 gallus gall
45	276	11.4	389	13 Q9W643	Q9w643 gallus gall

ALIGNMENTS

RESULT 1

Q9ULS1 PRELIMINARY; PRT; 532 AA.
ID Q9ULS1
AC Q9ULS1;
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-OCT-2001 (TREMREL. 18, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE KIAA1149 protein (fragment).
GN KIAA1149.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirose M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
RT "Characterization of cDNA clones selected by the Genemark analysis
RL DNA Res. 6:329-336(1999).
DR EMBL; AB032975; BAA86463.2;
DR HSSP; P56272; IAM5;
DR MEROPS; A01.004;
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 532 AA; 58720 MW; 98B135D0D5FBD2E8 CRC64;
Query Match 100.0%; Score 2419; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 1.7e-192;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy i ETDEPEPGRGRSGFVEMVDNLKSGQGYVEMTVGSPQTLNLVDVTGSSNFVGAAP 60
|||||
Db 77 ETDEPEPGRGRSGFVEMVDNLKSGQGYVEMTVGSPQTLNLVDVTGSSNFVGAAP 136

```
QY 61 HPFLHRYQRLSSTYRDLRKGVVYPYTGKWEGLGTLVSIHPGPNVTVRANIAAITE 120
|||||
Db 137 HPFLHRYQRLSSTYRDLRKGVVYPYTGKWEGLGTLVSIHPGPNVTVRANIAAITE 196
|||||
QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDLPPFDSLVKQTHVPNLFSLQLCGAGFPLNQS 180
|||||
Db 197 SDRFFINGSNWEGILGLAYAEIARPDSDLPPFDSLVKQTHVPNLFSLQLCGAGFPLNQS 256
|||||
QY 181 EVLASVGSMTIGGIDHSLYTGLSWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 240
|||||
Db 257 EVLASVGSMTIGGIDHSLYTGLSWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 316
|||||
QY 241 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVVCWQAGTTPWNIFPVIS 300
|||||
Db 317 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVVCWQAGTTPWNIFPVIS 376
|||||
QY 301 LYLMGEVNTGSRITILPQQYLRLPVEDVATSDDCYKFAISQSSTGCTVMGAVIMEGFYV 360
|||||
Db 377 LYLMGEVNTGSRITILPQQYLRLPVEDVATSDDCYKFAISQSSTGCTVMGAVIMEGFYV 436
|||||
QY 361 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
|||||
Db 437 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 496
|||||
QY 421 CALFMPLCLMVCWRCRLCRLRQHQHDDFADDISLLK 456
|||||
Db 497 CALFMPLCLMVCWRCRLCRLRQHQHDDFADDISLLK 532
|||||
```

RESULT 2

```
Q9BYC1
ID Q9BYC1 PRELIMINARY; PRT; 476 AA;
AC Q9BYC1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-site APP cleaving enzyme I-476.
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production."
RL Neurosci. Lett. 307:9-12(2001).
DR EMBL; AB050436; BAB40931.1; -.
DR HSSP; P32329; IYPS.
DR InterPro; IPR001461; AspproteaseA1.
DR DR PRINTS; PR00792; PRPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 476 AA; 52907 MW; 6C8C87F8A953AF66 CRC64;
```

Query Match 93.6%; Score 2264.5; DB 4; Length 476;
Best Local Similarity 94.5%; Pred. No. 9,9e-180;
Matches 431; Conservative 0; Mismatches 0; Indels 25; Gaps 1;

```
QY 1 ETDEEPEEPGRGSGFVEMVDNLRKSGGQYVYEMTVGSPPTNLILVDTGSSNFVAGAAP 60
|||||
Db 46 ETDEEPEEPGRGSGFVEMVDNLRKSGGQYVYEMTVGSPPTNLILVDTGSSNFVAGAAP 105
|||||
QY 61 HPFLHRYQRLSSTYRDLRKGVVYPYTGKWEGLGTLVSIHPGPNVTVRANIAAITE 120
|||||
Db 106 HPFLHRYQRLSSTYRDLRKGVVYPYTGKWEGLGTLVSIHPGPNVTVRANIAAITE 165
|||||
```

```
QY 121 SDRFFINGSNWEGILGLAYAEIARPDSDLPPFDSLVKQTHVPNLFSLQLCGAGFPLNQS 180
|||||
Db 166 SDRFFINGSNWEGILGLAYAEIARPDSDLPPFDSLVKQTHVPNLFSLQLCGAGFPLNQS 200
|||||
QY 181 EVLASVGSMTIGGIDHSLYTGLSWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 240
|||||
Db 201 EVLASVGSMTIGGIDHSLYTGLSWYTPIRREWYVEVILVRVEINGODLKMCKEYNDK 260
|||||
QY 241 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVVCWQAGTTPWNIFPVIS 300
|||||
Db 261 SIIVDSGTTNLRPKKVFEEAAVKSIAAASSTKFPDGFGLGEQLVVCWQAGTTPWNIFPVIS 320
|||||
QY 301 LYLMGEVNTGSRITILPQQYLRLPVEDVATSDDCYKFAISQSSTGCTVMGAVIMEGFYV 360
|||||
Db 321 LYLMGEVNTGSRITILPQQYLRLPVEDVATSDDCYKFAISQSSTGCTVMGAVIMEGFYV 380
|||||
QY 361 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 420
|||||
Db 381 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAI 440
|||||
QY 421 CALFMPLCLMVCWRCRLCRLRQHQHDDFADDISLLK 456
|||||
Db 441 CALFMPLCLMVCWRCRLCRLRQHQHDDFADDISLLK 476
|||||
```

RESULT 3

```
Q9BYC0
ID Q9BYC0 PRELIMINARY; PRT; 457 AA.
AC Q9BYC0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-site APP cleaving enzyme I-457 (Beta-site APP cleaving enzyme
DE type C).
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production."
RL Neurosci. Lett. 307:9-12(2001).
DR EMBL; AB050437; BAB40932.1; -.
DR HSSP; P32329; IYPS.
DR InterPro; IPR001461; AspproteaseA1.
DR DR PRINTS; PR00792; PRPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 457 AA; 51068 MW; C794C9A9E85FE7A2 CRC64;
```

Query Match 89.3%; Score 2161; DB 4; Length 457;
Best Local Similarity 90.4%; Pred. No. 3,7e-171;
Matches 412; Conservative 0; Mismatches 44; Gaps 1;

```
QY 1 ETDEEPEEPGRGSGFVEMVDNLRKSGGQYVYEMTVGSPPTNLILVDTGSSNFVAGAAP 60
|||||
Db 46 ETDEEPEEPGRGSGFVEMVDNLRKSGGQYVYEMTVGSPPTNLILVDTGSSNFVAGAAP 105
|||||
```



```
QY 191 IIGGIDSLTSGSLWYTPIRREWEYVEIIVRVEINGODLMDCKEYNDKSIYVDSGTNL 250
DB 1 IIGGIDSLTSGSLWYTPIRREWEYVEIIVRVEINGODLMDCKEYNDKSIYVDSGTNL 60
QY 251 RLPKKVFEAAVKSIIKASSTKFPDGEWLGEOVCWQAGTTWNIPFPVSLYLMGEVTNQ 310
DB 61 RLPKKVFEAAVKSIIKASSTKFPDGEWLGEOVCWQAGTTWNIPFPVSLYLMGEVTNQ 120
QY 311 SFRITILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGF 370
DB 121 SFRITILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGF 180
QY 371 AVSACHVHDEFRTAAVGGPFTVLMEDCGYNIPOQDESLMTIAYVMAAICAFMLPLCL 430
DB 181 AVSACHVHDEFRTAAVGGPFTVLMEDCGYNIPOQDESLMTIAYVMAAICAFMLPLCL 240
QY 431 MVCWRCLRLCRHQHDDFADDISLLK 456
DB 241 MVCWRCLRLCRHQHDDFADDISLLK 266

RESULT 6
Q9H2V8 PRELIMINARY; PRT; 439 AA.
AC Q9H2V8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDA13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PHOCHROMOCYTOMA;
RA Li Y., Huang Q., Peng, Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
RA Han Z.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212252; AAG41783.1; -
DR HSP; P00797; 2REN.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 439 AA; 48275 MW; 02EC0E050F11602 CRC64;

Query Match 48.0%; Score 1160.5; DB 4; Length 439;
Best Local Similarity 50.1%; Pred. No. 5e-88;
Matches 220; Conservative 78; Mismatches 134; Indels 7; Gaps 4;

QY 18 MVDNLRGKSGQYVVEVTGSPPTLTNLVDTGSSNFVCAAPHPFLHRYIQRLSTYR 77
DB 1 MVDNLRGKSGQYVVEVTGSPPTLTNLVDTGSSNFVCAAPHPFLHRYIQRLSTYR 77
QY 78 DLKRGYVVPYTGQKWEGLDLSIPHPNVTVRANIAAITESDKPFI 137
DB 61 SKGFDVTYKVTGQSWTGFVGEDLVTPKGFNTSFLVNIAITFESNFFLPKIKWNGILGL 120
QY 138 AYAEIARPDLSLEPFFDSLVKQTHVPLNLSLQCCAGFLNOSSEVLASVGGSMIGIDH 197
DB 121 AYATLAKPSSLSFTFFDSLVKQTHVPLNLSLQCCAGFLNOSSEVLASVGGSMIGIEP 177
QY 198 SLYTGSWYTPIRREWEYVEIIVRVEINGODLMDCKEYNDKSIYVDSGTNLRLPKVF 257
DB 178 SLYKGIWYTPIRREWEYVEIIVRVEINGODLMDCKEYNDKSIYVDSGTNLRLPKVF 237
QY 258 EAAVKSIKASSTKFPDGEWLGEOVCWQAGTTWNIPFPVSLYLMGEVTNQSFRTIL 317
DB 238 DAVVAVARASLIPEFSDGFWGTSQACWNTSETPWSYFPKISLYLRDENSRSRFTIL 297
QY 318 PQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHV 377
```

```
DB 298 PQYIQPMGAGLNY-ECYREGISPTNALVIGATVMEGFYVIFDRQKRVGFAASPCA 356
QY 378 HDEFRTAAVGGPFTVLMEDCGYNIPOQDESLMTIAYVMAAIC-ALFMLPLCLMVCOWR 436
DB 357 IAGAAVSEISGPFSTEDVASNCVPAQSLSEPIWIVSYALMSVCGAIIIVLIVLLPFR 416
QY 437 CLRLCRHQHDDFADDISLL 455
DB 417 CQR--RPRDPEVNDDESSL 433

RESULT 7
Q9JL18 PRELIMINARY; PRT; 514 AA.
AC Q9JL18;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartyl protease 1.
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Choi D.K., Sugano S., Sakaki Y.;
RT "Molecular characterization of the mouse Aspl gene, a homolog of the
human Aspl (Down Syndrome Region aspartyl protease).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216310; AAF36599.1; -
DR HSP; P32329; IYPS.
DR MEROPS; A01.041; -
DR MGD; MGI:1860440; Bace2.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 514 AA; 55799 MW; A70725F2C1DF5B47 CRC64;

Query Match 47.4%; Score 1146.5; DB 11; Length 514;
Best Local Similarity 50.7%; Pred. No. 9.3e-87;
Matches 216; Conservative 75; Mismatches 130; Indels 5; Gaps 3;

QY 8 EPCR-RGSFVEMVDNLRGKSGQYVVEVTGSPPTLTNLVDTGSSNFVCAAPHPFLHR 66
DB 65 EPVRATANFLAMVDNLRGKSGQYVVEVTGSPPTLTNLVDTGSSNFVCAAPHPFLHR 124
QY 67 YYQRSLSTYRDLRKGVPVYTGQKWEGLDLSIPHPNVTVRANIAAITESDKPFI 126
DB 125 YFDSSESSTYHSGFDVTYKVTGQSWTGFVGEDLVTPKGFNSFLVNIAITFESNFFL 184
QY 127 NGSNWEGILGAYAEIARPDLSLEPFFDSLVKQTHVPLNLSLQCCAGFLNOSSEVLASV 186
DB 185 PGIKWNGILGAYAAALAKPSSLSLETFDLSVAQAQKIPDIFSMOMCGAGLVAGS--CTN 241
QY 187 GGSMLIGIDHSLYTGSWYTPIRREWEYVEIIVRVEINGODLMDCKEYNDKSIYVDSG 246
DB 242 GGSVLGIGIPLSLYKGIWYTPIRREWEYVEIIVRVEINGODLMDCKEYNDKSIYVDSG 301
QY 247 TTNLRLPKVFEEAAVKSIKASSTKFPDGEWLGEOVCWQAGTTWNIPFPVSLYLMGE 306
DB 302 TTNLRLPKVFEEAAVKSIKASSTKFPDGEWLGEOVCWQAGTTWNIPFPVSLYLMGE 361
QY 307 VTNQSFRITILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARK 366
DB 362 NASRSFRITILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARK 420
QY 367 RIGFAVSACHVHDEFRTAAVGGPFTVLMEDCGYNIPOQDESLMTIAYVMAAICAFML 426
```

```
Db 421 RVGFAVSPCAIEGTGTTVSEISGPFSTEDIASNCVPAQALNEPILWIVSVALMSVCGAII 480
QY 427 PCLMW 432
Db 481 VLILL 486

RESULT 8
Q9NZL2 PRELIMINARY; PRT; 468 AA.
AC Q9NZL2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RL Cytochrome. Cell Genet. 89:177-184(2000).
DR EMBL: AF188276; AAF35835.1; -
DR HSP: P00797; 2REN.
DR MEROPS: A01.041; -
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 468 AA; 50324 MW; 717E0920126A0142 CRC64;

Query Match 40.2%; Score 972.5; DB 4; Length 468;
Best Local Similarity 43.1%; Pred. No. 2.3e-72;
Matches 193; Conservative 74; Mismatches 124; Indels 57; Gaps 5;

QY 9 PGRGSEFVEMVDNLKSGOGYVVTGSPPTNLILVDTGSSNFVAGAAPHPFLHRY 68
Db 71 PAGAANFLAMVDNLQDSGRGYYLELIGTPPKQLQLVDTGSSNFVAGTPHSYDITYF 130
QY 69 QRLSTYDLKRGVYVPTQGWEGELGTDLVSIPIHGNVTVRANIAATESDKFFING 128
Db 131 DTERSSTYKSGFDVTYKYGTSWTGFGEDLVTPKGFNTSLVNIATIFESENFPLP 190
QY 129 SNNEGILGLAYAEIARPDSDLEPFDLSLVKOTHPNLFSLQCGAGFPLNOSEVLASVG 188
Db 131 DTERSSTYKSGFDVTYKYGTSWTGFGEDLVTPKGFNTSLVNIATIFESENFPLP 190
QY 129 SNNEGILGLAYAEIARPDSDLEPFDLSLVKOTHPNLFSLQCGAGFPLNOSEVLASVG 188
Db 191 IKWNGILGLAYATLAKPSSLETFDLSLVTOANIPNVFSMQMGAGLPVAGS---GTNGG 247
QY 189 SMITGGIDHSLYGLSWYTPIRREWEYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTT 248
Db 248 SLVGGIEPSLYKGDWYTPKEEYQIEILKLEIGGOSLNDREYNADKAIVDSGTT 307
QY 249 NLRPLKPVFAAVKSKAASSTKFPDGFGLGQVLCWQAGTTPWNIFFVISLYLMGEVT 308
Db 248 SLVGGIEPSLYKGDWYTPKEEYQIEILKLEIGGOSLNDREYNADKAIVDSGTT 307
QY 249 NLRPLKPVFAAVKSKAASSTKFPDGFGLGQVLCWQAGTTPWNIFFVISLYLMGEVT 308
Db 308 LLRLPKVFDVAVVEAVARASLLPEFSDGFWTGSOLACWTNSETPWSVFPKISIVLRDENS 367
QY 309 NQSFRTILPQOYLRPVEDVATSDQDCYKFAISQSGTGTVMGAVINEGFFVDFDRAKRI 368
Db 308 LLRLPKVFDVAVVEAVARASLLPEFSDGFWTGSOLACWTNSETPWSVFPKISIVLRDENS 329
QY 309 NQSFRTILPQOYLRPVEDVATSDQDCYKFAISQSGTGTVMGAVINEGFFVDFDRAKRI 368
Db 330 -----YIQPMMGAGLNY-ECYREGISPTNALVIGATVMGEGFYIFDRAQKRV 376
QY 369 GFAVSACHVDEPRTAAVEGPTFLDMEGCGYNIPTDDESTLMTIAVMAAIC-ALFMPLP 427
Db 377 GFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPIILWIVSVALMSVCGAII 436
QY 428 LCLMVCWRCRLCRQHQHDDFADDISLL 455
Db 437 IVLLLLPFCOR--RPRDPEVNDDESL 462
```

```
RESULT 9
Q9NZL1 PRELIMINARY; PRT; 396 AA.
AC Q9NZL1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RL Cytochrome. Cell Genet. 89:177-184(2000).
DR EMBL: AF188277; AAF35836.1; -
DR HSP: P00797; 2REN.
DR MEROPS: A01.041; -
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 396 AA; 43013 MW; 5023A7AF391CEAC9 CRC64;

Query Match 40.0%; Score 967.5; DB 4; Length 396;
Best Local Similarity 54.6%; Pred. No. 4.7e-72;
Matches 183; Conservative 54; Mismatches 85; Indels 13; Gaps 4;

QY 9 PGRGSEFVEMVDNLKSGOGYVVTGSPPTNLILVDTGSSNFVAGAAPHPFLHRY 68
Db 71 PAGAANFLAMVDNLQDSGRGYYLELIGTPPKQLQLVDTGSSNFVAGTPHSYDITYF 130
QY 69 QRLSTYDLKRGVYVPTQGWEGELGTDLVSIPIHGNVTVRANIAATESDKFFING 128
Db 131 DTERSSTYKSGFDVTYKYGTSWTGFGEDLVTPKGFNTSLVNIATIFESENFPLP 190
QY 129 SNNEGILGLAYAEIARPDSDLEPFDLSLVKOTHPNLFSLQCGAGFPLNOSEVLASVG 188
Db 191 IKWNGILGLAYATLAKPSSLETFDLSLVTOANIPNVFSMQMGAGLPVAGS---GTNGG 247
QY 189 SMITGGIDHSLYGLSWYTPIRREWEYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTT 248
Db 248 SLVGGIEPSLYKGDWYTPKEEYQIEILKLEIGGOSLNDREYNADKAIVDSGTT 307
QY 249 NLRPLKPVFAAVKSKAASSTKFPDGFGLGQVLCWQAGTTPWNIFFVISLYLMGEVT 308
Db 308 LLRLPKVFDVAVVEAVARASLLPEFSDGFWTGSOLACWTNSETPWSVFPKISIVLRDENS 367
QY 309 NQSFRTILPQOYLRPVEDVATSDQDCYKFAISQ 342
Db 368 SRSFRTILPQK-LRVLQ-----CLKFPGLSQ 393

RESULT 10
Q9P0D2 PRELIMINARY; PRT; 213 AA.
AC Q9P0D2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE HSPC104 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CORD BLOOD;
RA Zhang Q.H., Ye M., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
RT "Human partial CDS cloned from cd34+ stem cells.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF161367; AAF28927.1; -.
DR InterPro: IPR001461; AspproteaseA1.
DR Pfam: PF00026; asp; 1.
FT NON_TER 1 1
SQ SEQUENCE 213 AA; 24338 MW; EC9D3FA31CFA835C CRC64;

Query Match 29.5%; Score 712.5; DB 4; Length 213;
Best Local Similarity 83.5%; Pred. No. 2.9e-51;
Matches 137; Conservative 4; Mismatches 12; Indels 11; Gaps 1;

QY 193 GGIDHSLYTGSWMYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLR 252
DB 1 GGIDHSLYTGSWMYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLR 60

QY 253 PKVFEAAVKSIRAASTKRPDPGFWLGEQLVCWQAGTTPWNIFFVISLYLMEVNTNOSF 312
DB 61 PKVFEAAVKSIRAASTKRPDPGFWLGEQLVCWQAGTTPWNIFFVISLYLMEVNTNOSF 120

QY 313 RITILPQQYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEG 356
DB 121 RITILPQQYLPR-----WKWPRPKTIVTVCHLTIVNG 153

RESULT 11
Q9R1P7
ID Q9R1P7 PRELIMINARY; PRT; 255 AA.
AC Q9R1P7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Aspartyl protease (Fragment).
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Accarino M., Fumagalli P., Taramelli R., Ottolenghi S.;
RT "Cloning of a gene from chromosome 21 Down Region encoding a potential
RT transmembrane protease.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF051150; AAD45964.1; -.
DR MEROPS: A01.041; -.
DR MGD: MGI:1860440; Bace2.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR PRINTS: PRO0792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
KW Protease.
FT NON_TER 1 1
SQ SEQUENCE 255 AA; 28685 MW; 53DE317815996D63 CRC64;

Query Match 24.7%; Score 596.5; DB 11; Length 255;
Best Local Similarity 47.8%; Pred. No. 1.7e-41;
Matches 109; Conservative 44; Mismatches 74; Indels 1; Gaps 1;

QY 205 WYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKKVEAAVKSI 264
DB 1 WYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKKVEAAVKSI 60

QY 265 KAASSTKFPDGEWLGQVLCWQAGTTPWNIFFVISLYLMEVNTNOSFRITILPQQYLPR 324
DB 61 ARTSLIPESDGEWLGQVLCWQAGTTPWNIFFVISLYLMEVNTNOSFRITILPQQYLPR 120
```

```
QY 325 VEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYYVVFDRARKRIGFAVSACHVHDEPFTA 384
DB 121 VMGAGFNY-ECYRFGISSSTNALVIGATVMEGYYVVFDRARRVGVFAVSCAEIEGTIVS 179

QY 385 AVGGPFVTLMDKCGYNIPQDTESLMTIAYVMAATCALPMLPLCLMV 432
DB 180 EISGPFSTEDIASNCVPAQALNEPLTWISYALMSVCGAILLVILL 227

RESULT 12
Q8WQY9
ID Q8WQY9 PRELIMINARY; PRT; 244 AA.
AC Q8WQY9
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartate protease (Fragment).
GN APP.
OS Aphrocallistes vastus.
OC Eukaryota; Metazoa; Porifera; Hexactinellida; Hexasterophora;
OC Lyssacinosida; Rosellidae; Aphrocallistes.
OX NCBI_TaxID=83887;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller W.E.G., Mueller I.M., Grebenjuk V.A.;
RT "Urmetazoa: Origin and evolution of the common ancestor of Metazoa.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ304863; CAC83293.1; -.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp; 1.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW Protease.
FT NON_TER 244 244
SQ SEQUENCE 244 AA; 26366 MW; 6536902661E0E4C7 CRC64;

Query Match 14.9%; Score 359.5; DB 5; Length 244;
Best Local Similarity 36.0%; Pred. No. 7.9e-22;
Matches 81; Conservative 41; Mismatches 90; Indels 13; Gaps 6;

QY 19 VDNLRKSGGGYVYEMTVGSPPTLNILVDTGSSNFAGVAAPHPFHRY--YORQLSSTY 76
DB 22 VYQLQGPESGYLLSVNLGTPPEKFLVLDGSSNFAVAAGFYIDYPRFDKSLSKTF 81

QY 77 RDLRKGVYVYPYTOCKWEGELGTDLVS----IPHGPNVTYRANIAAI--TESDKPFINGSNW 131
DB 82 RDINSEVGKVIIDGWSGRVGEDYFAFASDVTTNASKSVKVVYVSLIESVSEGFETSGGW 141

QY 132 EGIPLGAYABRIARPDDSLERPFDSLVKQTHVPNLFSLQLCGAGFPPLNQSE-VLASVCGSM 190
DB 142 VGLTGMGYAVLAKPDSSITPVMDLSLVSGVTSKDKGLQLCQ---PLSNSRELNVQNGKM 198

QY 191 IIGGDHSLYTGSS--LWYTPIRREWYEVIIVRVEINGQDLKMDCK 233
DB 199 SLGSARPLPSNSDVFRYVAITEESFYEVILTNIKVGSTSLDLP 243

RESULT 13
Q9VKP6
ID Q9VKP6 PRELIMINARY; PRT; 391 AA.
AC Q9VKP6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG17134 protein (R02351p).
GN CG17134.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=BERKELEY.
RX  MEDLINE=20196006; PubMed=107311132;
RA  Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borkov D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA  Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de
RA  Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimbel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Pacieb J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svrtkas R., Rector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
RT  "The genome sequence of Drosophila melanogaster."
RL  Science 287:2185-2195(2000).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN=BERKELEY.
RA  Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA  Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA  George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA  Miranda A., Mungall C.J., Nunoo J., Pacieb J., Paragas V., Park S.,
RA  Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA  Celniker S.;
RL  Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC  !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC  EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR  EMBL: AE003630; AAF53016.1; -.
DR  EMBL: AY070911; AAL48533.1; -.
DR  HSSP: P00794; 4CMS.
DR  FlyBase: FBgn0032304; CG17134.
DR  InterPro: IPR001461; AspproteaseA1.
DR  InterPro: IPR001969; Aspprotease_site.
DR  Pfam: PF00026; asp.1.
DR  PRINTS: PR00792; PEPSIN.
DR  PROSITE: PS00141; ASP_PROTEASE; 2.
KW  Aspartyl protease; Hydrolase.
SQ  SEQUENCE 391 AA; 42016 MW; 99A18E7131025E1B CRC64;

Query Match 13.7%; Score 332.5; DB 5; Length 391;
Best Local Similarity 29.1%; Pred. No. 2.6e-19;
Matches 105; Conservative 61; Mismatches 146; Indels 49; Gaps 13;

QY 20 DNLGRSGGQGYVEMTVGSPPTLNILVDGTSSNFAVGAAPHF-----LHRYQRQLS 73
Db 66 ENLHSMNNNEYGVIAIGTQQRNFIILFDGTGSANLWPSACPSATNCQRHNKYDSSAS 125

```

```

QY 74 STYRDLRKGVVVPYTOGKWEGLGTLVLSIPHGNVTVVRANIAAITEKDFKFFINGSNWEG 133
Db 126 STYVANGEEFAIEYGTGSLGFLSNDIVTIA-GISIQNTQFGEALSPEGTFVD-APFAG 183
QY 134 ILGLAVAEIARPDSDLEPFDSLVKQHPVN-LFSLQCGAGPPLNOSEVLASVGSMMII 192
Db 184 ILGLAFSAIA--VDGVTPTPPDNMISQGLDEPVSFLKRGQ-----TAVRGEGIL 233
QY 193 GGDHSLYTSGLMYTPIRREWYVEIIVRVEINGODLKMCKEYNIKSDIVDSGTTNLR 252
Db 234 GGIDSSLYRSLTVVPVSPAYWQFKVNTIKTNGTLLCNGC-----QALADTGTSLIAV 287
QY 253 PKVFFAAVKSIAKASSTEREPDGFGLGQLV-CWQAGTTPWNIFPVISLYLMEVNTQS 311
Db 288 PLAYRKINQLGATDND-----GEAFVRCGRVSS-----LPKVNINIGTV---- 329
QY 312 FRITLPQOYLRPVEDVATSDQCYK-FAISQSTGTVMGAVINEGVYVDFDRARKRIGF 370
Db 330 --FTLAPRDYI--VKVTPQNGQTYCMSAFTYMEGLSFLGDFVFGTGYTFVDKGNRIGF 385
QY 371 A 371
Db 386 A 386

RESULT 14
Q9GYX7 PRELIMINARY; PRT; 354 AA.
ID Q9GYX7 PRELIMINARY; PRT; 354 AA.
AC Q9GYX7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Heme-binding aspartic proteinase (Fragment).
OS Boophilus microplus (cattle tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.
OX NCBI_TaxID=6941;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PORTO ALEGRE; TISSUE=OVARY;
RA Sorgine M.H.F., Logullo C., Zingali R.B., Paiva-Silva G.O.,
RA Juliano L., Oliveira P.L.;
RT "A heme-binding aspartic proteinase from the eggs of the hard tick
RT Boophilus microplus."
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL: AF286865; AAG00993.1; -.
DR HSSP: P00797; 2REN.
DR MEROPS: A01.054; -.
DR InterPro: IPR001461; AspproteaseA1.
DR Pfam: PF00026; asp.1.
DR PRINTS: PR00792; PEPSIN.
DR NON_TER 1
FT SEQUENCE 354 AA; 38512 MW; B50D7C638CF27091 CRC64;

Query Match 13.7%; Score 332; DB 5; Length 354;
Best Local Similarity 26.2%; Pred. No. 2.6e-19;
Matches 101; Conservative 74; Mismatches 135; Indels 76; Gaps 15;

QY 19 VDNLRGK-----SGQGYVEMTVGSPPTLNILVDGTSSNFAVGAAPHFPL- 64
Db 13 VTEIRGALGDPPIILNTNMQFYGIITGTPQSFKLMDTGSNFWFJ---PSIN 67
QY 65 -----HRYQRQLSSTYRDLKGVVVPYTOGKWEGLGTLVLSIPHGNVTVVRANI 115
Db 68 CDOSMACRDHAKYDSSKSTFTKSGRYIRIRYSGGVVGRGITSIDNVGV--GPATVVTQYKF 125
QY 116 AATITESDKFFINGSNWEGILGLVAETARPDSDLEPFDSLVKQHPVN-LFSLQCGAG 174
Db 126 AEMHSDGKLFNRNAKYDITGLAPPSISQ--NNOLPLFDAMVKGVVRAQVAFSLYL--SK 181
QY 175 FPLNQSEVLASVGSMMIIGGIDHSLYTSGLMYTPIRREWYVEIIVRVEINGODLKM-DC 233

```

Db 182 QPSEQN-----GGEIYFGGINAQRYTGAIHVVPVSQAHHQVVMNDINVGQTTLCVGGC 235
QY 234 KEYNDSKSIIVSDSGTNNLRPKKVFEEAAVKSIAKAASSTKFPDGFGEQLVCMQAGTTPW 293
Db 236 -----PTVDSGTSFSLGP-----SADVETLNRVIGATTAAGY-----FEVNCATI 277
QY 294 NIFPVISLYLME---VTNQSFRTILPQQYLRLPVEDVATSDQDCYKFAISQSTGT--- 347
Db 278 SSLPPTIFNLNGKSFPLQGEAYTIRI-----PLTTGGQCQCFTRISESDASGTNLW 327
QY 348 VMGAVIMEGFYVDFDRARKRIGFAVS 373
Db 328 ILGAVETQTYTYTVFDRAGNRVGFATA 353

RESULT 15

Q9BGU5 PRELIMINARY; PRT: 386 AA.
AC Q9BGU5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cathepsin D (Fragment).
GN CAT-D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Higuchi M., Miyashita N., Nagamine Y., Awata T.;
RT "Complementary DNA sequence of bovine cathepsin D.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DDI databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR EMBL; AB055312; BAB21620.1; -
DR HSSP; P07339; 1LYB.
DR MEROPS; A01.009; -
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp.1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Aspartyl protease; Hydrolase.
FT NON_TER 1
SQ SEQUENCE 386 AA; 41965 MW; 33BF00B080585490D CRC64;

Query Match 12.98; Score 312.5; DB 6; Length 386;
Best Local Similarity 28.08; Pred. NO. 1.3e-17;
Matches 113; Conservative 69; Mismatches 131; Indels 91; Gaps 20;
QY 8 EPG-RGGSFVEMVDNLRGKSGQGYVEMTVGSPPTNLNLDVTGSSNFVGAAPHPEL-- 64
Db 35 EPAVRQGPPELLKNYMDAQ---YGEIGTGPPOCFVTFVDFGSLANLWPSIHCULKLDI 91
QY 65 ----HRYYQRLSSTRDLRRGVY--VPYTGKWEGLGTLVLSIPHGN-----VTVR 112
Db 92 ACWTHRKYNSDKSSTY--VKNGTTFDIHYGSGSLGSLYLSQDTSVPCNPSSSPGGVTQ 149
QY 113 ANT--AAITSDKFFINGNWEGLGLAYAEIARDDSLPFDLSLVKQTHV-PNLFSLQ 169
Db 150 RQTFGEAIKQPGVVFI-AAKFGGLGMAYPRIS--VNNVLPVFDNLMMQKLVOKNVFS-- 204
QY 170 LCGAGPLNQSEVLASVGGSMILGIDHSLSYGLWYTPIRREWYVEIIVRVEINGDL 229
Db 205 -----FFLNR-DPRAQPGGELMLGTDTSKYRGSLMFHNVTRQAYWOIHMDQLDV-GSSL 257
QY 230 KMDCKEYNYSKSIIVSDSGTNNLRPKKVFEEAAVKSIAKAASSTKFPDGFGEQLVCMQAG 289
Db 258 TV-CK--GGCEAIVDTGTSLVGVPVEEVRLEQAI-----G 290
QY 290 TTPWNIFPVISLYLMGEVNTNOSFRITILPQQYLR-PVEDVATSDQDCYKFAISQSTGT- 347

Db 291 AVP-----LIQGEYMIPCVKVSSLPQVTVKLGKDYAXSPED-YALKVSOAGTTCVC 340
QY 348 -----VMGAVIMEGFYVDFDRARKRIGFAVS 374
Db 341 LSGFMGMDIPPPGGLWILGDVFIGRYTYTVFDRQNRVGLAEAA 384

Search completed: March 4, 2003, 10:04:29
Job time : 40 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2003, 13:38:25 ; Search time 4002 Seconds
(without alignments)
17074.810 Million cell updates/sec

Title: US-09-723-722A-44
Perfect score: 2348
Sequence: 1 ccacgcggccctcacagc.....caagattgcctcttgatt 2348

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl : *

1: gb_ba : *

2: gb_htg : *

3: gb_in : *

4: gb_om : *

5: gb_ov : *

6: gb_pat : *

7: gb_ph : *

8: gb_pl : *

9: gb_pr : *

10: gb_ro : *

11: gb_sts : *

12: gb_sy : *

13: gb_un : *

14: gb_vi : *

15: em_ba : *

16: em_fun : *

17: em_hum : *

18: em_in : *

19: em_mu : *

20: em_om : *

21: em_or : *

22: em_ov : *

23: em_pat : *

24: em_ph : *

25: em_pl : *

26: em_ro : *

27: em_sts : *

28: em_un : *

29: em_vi : *

30: em_htg_hum : *

31: em_htg_inv : *

32: em_htg_other : *

33: em_htg_mus : *

34: em_htg_pln : *

35: em_htg_rod : *

36: em_htg_mam : *

37: em_htg_vrt : *

38: em_sy : *

39: em_htgo_hum : *

40: em_htgo_mus : *

41: em_htgo_other : *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2342.8	99.8	2526	9	AF190725	AF190725 Homo sapi
2	2341.2	99.7	5878	9	AF201468	AF201468 Homo sapi
3	2331.8	99.3	5814	9	AB032975	AB032975 Homo sapi
4	2181.2	92.9	5757	6	AX364933	AX364933 Sequence
5	2041.4	86.9	2070	9	AF200343	AF200343 Homo sapi
6	2039.8	86.9	2070	6	AX105385	AX105385 Sequence
7	2039.8	86.9	2070	6	AX378015	AX378015 Sequence
8	2024	86.2	2541	6	AR178469	AR178469 Sequence
9	2024	86.2	2541	6	AX002655	AX002655 Sequence
10	2024	86.2	2541	6	E50816	E50816 Aspartate p
11	2004	85.3	3252	6	AX062111	AX062111 Sequence
12	2004	85.3	3252	6	AX063201	AX063201 Sequence
13	2004	85.3	3252	6	AX472368	AX472368 Sequence
14	2004	85.3	3252	9	AF200193	AF200193 Homo sapi
15	1881	80.1	1977	6	AX105387	AX105387 Sequence
16	1881	80.1	1977	6	AX378017	AX378017 Sequence
17	1835.4	78.2	2370	6	AR178470	AR178470 Sequence
18	1835.4	78.2	2370	6	AX002657	AX002657 Sequence
19	1835.4	78.2	2370	6	E50817	E50817 Aspartate p
20	1660.6	70.7	3880	10	AF190726	AF190726 Mus muscu
21	1550	66.0	2043	6	AX105389	AX105389 Sequence
22	1550	66.0	2043	10	AF200346	AF200346 Mus muscu
23	1511.2	64.4	2158	6	AX401989	AX401989 Sequence
24	1511.2	64.4	2158	10	AF190727	AF190727 Rattus no
25	1506	64.1	1506	9	AF204943	AF204943 Homo sapi
26	1380	58.8	1485	9	AB050436	AB050436 Homo sapi
27	1355.8	57.7	1362	6	AX105411	AX105411 Sequence
28	1355.8	57.7	1380	6	AX105413	AX105413 Sequence
29	1344.4	57.3	1431	9	AF338816	AF338816 Homo sapi
30	1298	55.3	1506	6	AR202196	AR202196 Sequence
31	1293.8	55.1	1341	6	AX105403	AX105403 Sequence
32	1293.8	55.1	1380	6	AX105405	AX105405 Sequence
33	1270.4	54.1	1302	6	AX105407	AX105407 Sequence
34	1266	53.9	1408	9	AB050437	AB050437 Homo sapi
35	1230.4	52.4	1374	9	AF338817	AF338817 Homo sapi
36	1197.4	51.0	1287	6	AX105432	AX105432 Sequence
37	1197.4	51.0	1305	6	AX105434	AX105434 Sequence
38	1180	50.3	1278	6	AX105409	AX105409 Sequence
39	1116	47.5	1333	9	AB050438	AB050438 Homo sapi
40	990	42.2	1114	9	AF161367	AF161367 Homo sapi
C 41	781.2	33.3	134278	9	AP001822	AP001822 Homo sapi
C 42	781.2	33.3	199892	2	AC020997	AC020997 Homo sapi
C 43	779.6	33.2	98305	2	AP000685	AP000685 Homo sapi
C 44	779.6	33.2	149843	9	AP000892	AP000892 Homo sapi
C 45	779.6	33.2	162610	2	AP000761	AP000761 Homo sapi

ALIGNMENTS

RESULT 1

AF190725

LOCUS

DEFINITION Homo sapiens beta-site APP cleaving enzyme (BACE) mRNA, complete cds.

ACCESSION AF190725

VERSION AF190725.1

KEYWORDS GI:6118538

SOURCE

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2526)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Vassar, R., Bennett, B.D., Babu-Khan, S., Kahn, S., Mendiaz, E.A., Denis, P., Teplow, D.B., Ross, S., Amarante, P., Loeloff, R., Luo, Y.,

Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M.A.,
Biern, A.L., Curran, E., Burgess, T., Louis, J.C., Collins, F.,
Treanor, J., Rogers, G., and Citron, M.
Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
the transmembrane aspartic protease BACE

JOURNAL
MEDLINE
PUBMED
20002972
10531052

REFERENCE

2 (bases 1 to 2526)

Bennett, B.D., Vassar, R., and Citron, M.

Direct Submission

Submitted (29-SEP-1999) Neuroscience, Amgen Inc., One Amgen Center

Dr., Thousand Oaks, CA 91320-1799, USA

Location/Qualifiers

FEATURES

1..2526

/organism="Homo sapiens"

/db_xref="taxon:9606"

1..2526

/gene="BACE"

454..1959

/gene="BACE"

/codon_start=1

/product="beta-site APP cleaving enzyme"

/protein_id="AAF04142.1"

/db_xref="GI:6118539"

/translation="NAQALPWLMLWMGAGVLPAGHTQHGIRLPLRSLGLGAPLGLRLP

RETDEPEPEGRSGFVEMVDNLKRGSGGYVYEMVGPSPOTNLIVDTGSSNFAVG

AAPHELFHYQROLSTVYDLKRGVYVPTQCKWEGELGTDLVSIPIHGNVTVRANI

AATPELFFNGSNWGLGLAYELIARPDLSLEPFEDSLVKQTHVPLNLSLQCGA

GFPLNQSLVAGSMITGGIDHSITGLWPTIRREMYEIIIVREINGQDLKM

DKENYDKSIVDSGTTNLRPKKFEAAVSIKASSTSEKFPDGLVQWQAG

TTWNIFPVISLMEVTNQSFRIITLPQYLRPVEDVATISQDDCYKFAISQSSTGT

VMGAVIMEGFYVDFARRKRGIFAVSACHVDEFRTAAVEGPFVTLDMEDCGYNIPQT

DESTLMTAYVMAALCALFMLPLCLMLVCQWRCLRLRQHQHDFADDDISLLK"

BASE COUNT 534 a 781 c 715 g 496 t

ORIGIN

Query Match 99.8%; Score 2342.8; DB 9; Length 2526;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2344; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCATGCCGCGCCCTCACAGCCGCGCGGAGCCGCGCTGCCAGGCTGCGCGCGG 60
DB 149 CCATGCCGCGCCCTCCAGCCGCGCGGAGCCGCGCTGCCAGGCTGCGCGCGG 208
QY 61 CSGTGCCGATGTAGCGGCTCGCGGATCCAGCCCTCTCCCTGCTCCGCTGCTCGGAT 120
DB 209 CCGTGCCGATGTAGCGGCTCGCGATCCAGCCCTCTCCCTGCTCCGCTGCTCGGAT 268
QY 121 CTCCCTGACCGCTTCCACAGCCGAGCCGCGGCTGCGCCAGGCGCTGCGAGCCCT 180
DB 269 CTCCCTGACCGCTTCCACAGCCGAGCCGCGGCTGCGCCAGGCGCTGCGAGCCCT 328
QY 181 GGCTCTCTGATGCCCGCCAGCTCCCTCTCTCTGAGAGCCAGCCAGCCAGACTGG 240
DB 329 GGCTCTCTGATGCCCGCCAGCTCCCTCTCTCTGAGAGCCAGCCAGACTGG 388
QY 241 GGGCAGGCGCCAGGAGCAGCTGGGCCAGTGGCCAGCCAGAGGCGCCGAAGCGCGGCG 300
DB 389 GGGCAGGCGCCAGGAGCAGCTGGGCCAGTGGCCAGCCAGAGGCGCCGAAGCGCGGCG 448
QY 301 CCACATGGCCCAAGCCCTGCGCTGCTGCTGTGGATGGGCGCGGAGTGTGCTGCTG 360
DB 449 CCACATGGCCCAAGCCCTGCGCTGCTGCTGTGGATGGGCGCGGAGTGTGCTGCTG 508
QY 361 CCCAGGCGCCAGCAGCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 509 CCCAGGCGCCAGCAGCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 568
QY 421 TGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 569 TGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528

```
|||||
Db 1709 CTTGCCATGTGCACGATGAGTTTCAGGACGGCGGTGAAGGCCCTTTTCTCACCTTGG 1768
Qy 1621 ACATGGAAGACTGTGGCTACACATTCACACAGACAGATGAGTCAACCTCATGACCATAG 1680
Db 1769 ACATGGAAGACTGTGGCTACACATTCACACAGACAGATGAGTCAACCTCATGACCATAG 1828
Qy 1681 CCTATGTCATGGCTGCCATCTGCCCTCTTCATGCTGCCACTCTGCTCATGTGTGTC 1740
Db 1829 CCTATGTCATGGCTGCCATCTGCCCTCTTCATGCTGCCACTCTGCTCATGTGTGTC 1888
Qy 1741 AGTGGCCTGCTCCGCTGCTGCGCCAGCAGCATGATGATTTGCTGATGACATCTCCC 1800
Db 1889 AGTGGCCTGCTCCGCTGCTGCGCCAGCAGCATGATGATTTGCTGATGACATCTCCC 1948
Qy 1801 TGCTGAAGTGAAGAGGCCATGGCCAGACAGATAGAGATTCCTCCCTGGACACACCTCCGTC 1860
Db 1949 TGCTGAAGTGAAGAGGCCATGGCCAGACAGATAGAGATTCCTCCCTGGACACACCTCCGTC 2008
Qy 1861 GTTCACCTTTGGTCAACAAGTAGGAGACACAGATGGCAGCTGTGGCCAGACACCTCAGGAC 1920
Db 2009 GTTCACCTTTGGTCAACAAGTAGGAGACACAGATGGCAGCTGTGGCCAGACACCTCAGGAC 2068
Qy 1921 CCTCCCAACCCACCAATGCTCTGCTTGATGGAGAGAAAGGCTGGCAAGGTGGT 1980
Db 2069 CCTCCCAACCCACCAATGCTCTGCTTGATGGAGAGAAAGGCTGGCAAGGTGGT 2128
Qy 1981 TCCAGGACTGTACCTGTAGGAACACAGAAAGAGAAAGAACACACTCTGCTGGCGGGA 2040
Db 2129 TCCAGGACTGTACCTGTAGGAACACAGAAAGAGAAAGAACACACTCTGCTGGCGGGA 2188
Qy 2041 ATACTCTTGGTCACTCAAAATTTAAGTCGGGAAATTCCTGCTGTTGAACTTCAGCCCTG 2100
Db 2189 ATACTCTTGGTCACTCAAAATTTAAGTCGGGAAATTCCTGCTGTTGAACTTCAGCCCTG 2248
Qy 2101 AACCTTTGTCACCATTCCTTTAAATTTCTCAACCCCAAGATTTCTTTCTTTAGTTT 2160
Db 2249 AACCTTTGTCACCATTCCTTTAAATTTCTCAACCCCAAGATTTCTTTCTTTAGTTT 2308
Qy 2161 CAGAAGTACTGGCATCACAGCGATTACCTTGGCGTGTGCTGTTGGTACCTGGCAG 2220
Db 2309 CAGAAGTACTGGCATCACAGCGATTACCTTGGCGTGTGCTGTTGGTACCTGGCAG 2368
Qy 2221 AGAAGACACCAAGCTTGTTCCTGCTGGCCAAAGTCACTAGGAGAGGATCACAGTTTG 2280
Db 2369 AGAAGACCAAGCTTGTTCCTGCTGGCCAAAGTCACTAGGAGAGGATCACAGTTTG 2428
Qy 2281 CTATTTCCTTTAGACAGGAGCTGTATTAACCAAGCCCTAACATTTGGTGTCAAGATTTGCT 2340
Db 2429 CTATTTCCTTTAGACAGGAGCTGTATTAACCAAGCCCTAACATTTGGTGTCAAGATTTGCT 2488
Qy 2341 CTTGAATT 2348
Db 2489 CTTGAATT 2496

RESULT 2
AF201468
LOCUS AF201468 5878 bp mRNA linear PRI 19-DEC-1999
DEFINITION Homo sapiens APP beta-secretase mRNA, complete cds.
ACCESSION AF201468
VERSION AF201468.1 GI:6601444
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5878)
Sinha, S., Anderson, J.P., Barbour, R., Basi, G.S., Caccavello, R.,
Davis, D., Doan, M., Dovey, H.F., Frigon, N., Hong, J.,
Jacobson-Croak, K., Jewett, N., Keim, P., Knops, J., Lieberburg, I.,
Power, M., Tan, H., Tatsuno, G., Tung, J., Schenk, D., Seubert, P.,
Suomensaari, S., Wang, S., Walker, D., Zhao, J., McConlogue, L. and
```

```
John, V.
Purification and cloning of amyloid precursor protein
beta-secretase from human brain
Nature 402 (6761), 537-540 (1999)
20051711
PUBMED 10591214
REFERENCE 2 (bases 1 to 5878)
AUTHORS Basi, G.S., Power, M.D., Wang, S., Tatsuno, G., Frigon, N., Doan, M.,
Hong, G., Keim, P., Anderson, J., Sinha, S. and McConlogue, L.M.
Direct Submission
Submitted (03-NOV-1999) Gene Expression Group, Elan
Pharmaceuticals, Inc., 800 Gateway Blvd., S. San Francisco, CA
94080, USA
FEATURES
Location/Qualifiers
1..5878
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="D11S4352"
/tissue_type="brain"
455..1960
/function="cleaves APP at the beta-cleavage site"
/note="membrane type aspartyl protease"
/codon_start=1
/product="APP beta-secretase"
/protein_id="AAPI8982.1"
/db_xref="GI:6601445"
/translation="MAQALPWLMLMGAGVLPAGHTQHGIRLPLRSGLGAPLGLRLP
REIDPEEPGRRGSEFVMDNLRKSGQYVEMTVGSPQTUNILVLDVSSNFVAG
AAPEFLHRYQRLSSYDLRKGVVPTQGWKGEGLDLSVIPHGNVTVRANI
AATESDKFPIINGSNMELIGLAYAEIARPDQSLPEFDSLKVHPNLFSLQCGA
GFPLNDSVLAISVSGMIIIGIDHSLVLTGSLWTPIRREWYEVIIYRVETNGDLAK
DCKEYANDKSIIVDSGTTNLRPKKVFEEAVKSIKAASSTKFPDGLVQVQWAG
TTWNIPFVISLYLMGEVNTQSFRTILPOQYLRPVEDVATSDQCKFKPAISQSTGT
VMGAVMEGFYVDFDRARRKIRFAVSHVIDEPTAAVEGPFVFLDMEDCGYNIPOT
DESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRQHQHDFADDISLLK"
BASE COUNT 1547 a 1518 c 1405 g 1408 t
ORIGIN

Query Match 99.7%; Score 2341.2; DB 9; Length 5878;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2343; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CCATGCGCGCCCTCACAGCCCGCGGAGCCGCGCTGCCAGCTGCCAGCTGGCGCGCG 60
Db 150 CCATGCGCGCCCTCCACAGCCCGCGGAGCCGCGCGCTGCCAGCTGGCGCGCG 209
Qy 61 CSGTGGCGATAGCGGCTCCGGATCCAGCCTCTCCCTGCTCCCTGCTCTGCGGAT 120
Db 210 CCGTGGCGATAGCGGCTCCGGATCCAGCCTCTCCCTGCTCTCCCTGCTCTGCGGAT 269
Qy 121 CTCCTGTCACCGCTCTCCACAGCCCGGAGCCGCGGCTGGCCAGGCCCTGCAGGCCCT 180
Db 270 CTCCTGTCACCGCTCTCCACAGCCCGGAGCCGCGGCTGGCCAGGCCCTGCAGGCCCT 329
Qy 181 GCGTCTCTGATGCCCCCAAGCTCTCTCTGAGAAGCCACAGCACCACCCAGACTTGG 240
Db 330 GCGTCTCTGATGCCCCCAAGCTCTCTCTGAGAAGCCACAGCACCACCCAGACTTGG 389
Qy 241 GGGCAGGCCCGCAGGACGGAGCTGGGCCAGTGCAGGCCAGAGGCCCGCGAGGCCGGGC 300
Db 390 GGGCAGGCCCGCAGGACGGAGCTGGGCCAGTGCAGGCCAGAGGCCCGCGAGGCCGGGC 449
Qy 301 CCACATGGCCCAAGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 450 CCACATGGCCCAAGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509
Qy 361 CCCAGGCCCGCAGCAGCGCATCCGGCTGCCCTCGCGCAGCGGCTGGGGCGGCCCGCCC 420
Db 510 CCCAGGCCCGCAGCAGCGCATCCGGCTGCCCTCGCGCAGCGGCTGGGGCGGCCCGCCC 569
Qy 421 TGGGCTCGCGCTGCGCCCGGGAGCCAGCAGAGAGCCCGAGAGCCCGCGCGCGAGGGGCA 480
```

Db 570 TGGGGCTGGGGCTGCCCCGGGAGACGACGAGCCGAGAGCCCGGCGGAGGGGCA 629
QY 481 GCTTTGTGGAGATGTTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGA 540
Db 630 GCTTTGTGGAGATGTTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGA 589
QY 541 TGACCGTGGGAGCCCCCGGAGACGCTCAACATCTCTGGTGGATACAGGACGACGTAAC 600
Db 690 TGACCGTGGGAGCCCCCGGAGACGCTCAACATCTCTGGTGGATACAGGACGACGTAAC 749
QY 601 TTGCAGTGGGTGCTGCCCCCCCCACCCCTTCTGCATCCTACTACAGAGGACGCTGTCCA 560
Db 750 TTGCAGTGGGTGCTGCCCCCCCCACCCCTTCTGCATCCTACTACAGAGGACGCTGTCCA 809
QY 661 GCACATACCGGGACCTCCCGAAGGTTGTATGTGCCCTACACCCAGGGCAAGTGGGAAG 720
Db 810 GCACATACCGGGACCTCCCGAAGGTTGTATGTGCCCTACACCCAGGGCAAGTGGGAAG 869
QY 721 GGGAGCTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCTGGCA 780
Db 870 GGGAGCTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCTGGCA 929
QY 781 ACATTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAAGGCTCCCAACTGGGAAGCA 840
Db 930 ACATTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAAGGCTCCCAACTGGGAAGCA 989
QY 841 TCCTGGGCTGGCCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTG 900
Db 990 TCCTGGGCTGGCCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCTTTCTTTG 1049
QY 901 ACTCTCTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCTCGAGCTTTGTGGTGTG 960
Db 1050 ACTCTCTGGTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCTCGAGCTTTGTGGTGTG 1109
QY 961 GCTTCCCTCAACCCAGTCTGAAGTCTGCGCTCTGTCGGAGGAGCATGATCATTTGGAG 1020
Db 1110 GCTTCCCTCAACCCAGTCTGAAGTCTGCGCTCTGTCGGAGGAGCATGATCATTTGGAG 1169
QY 1021 GTATCGACCTCGTGTACACAGGCGAGTCTCTGGTATACACCCATCCGGCGGAGTGT 1080
Db 1170 GTATCGACCTCGTGTACACAGGCGAGTCTCTGGTATACACCCATCCGGCGGAGTGT 1229
QY 1081 ATTATGAGTGTATGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGAGTGA 1140
Db 1230 ATTATGAGTGTATGTCGGGTGGAGATCAATGGACAGGATCTGAAATGGAGTGA 1289
QY 1141 AGGAGTCAACCTATCACAGAGCATTTGACAGTGGCACAGTGGCACCACTCTGTTGCCCA 1200
Db 1290 AGGAGTCAACCTATCACAGAGCATTTGACAGTGGCACAGTGGCACCACTCTGTTGCCCA 1349
QY 1201 AGAAAGTGTTCAGCTGCAAGTCAATCCATCAAGGAGGCTCTCCACGGAGAGTTC 1260
Db 1350 AGAAAGTGTTCAGCTGCAAGTCAATCCATCAAGGAGGCTCTCCACGGAGAGTTC 1409
QY 1261 CTGATGTTTCTGGGTAGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCACTTGGGA 1320
Db 1410 CTGATGTTTCTGGGTAGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCACTTGGGA 1469
QY 1321 ACATTTTCCAGTCACTACCTACCTAATGGTGGAGTTACCAACAGTCTTCCGCA 1380
Db 1470 ACATTTTCCAGTCACTACCTAATGGTGGAGTTACCAACAGTCTTCCGCA 1529
QY 1381 TCACCATCTTCCGAGCAATACCTGCGGCCAGTGGAAAGATGTGGCCACAGTCCCAAGAG 1440
Db 1530 TCACCATCTTCCGAGCAATACCTGCGGCCAGTGGAAAGATGTGGCCACAGTCCCAAGAG 1589
QY 1441 ACTGTTACAAGTTTCCCATCTCACAGTCACTCCAGGGCAGCTGTTATGGGAGCTGTATCA 1500
Db 1590 ACTGTTACAAGTTTCCCATCTCACAGTCACTCCAGGGCAGCTGTTATGGGAGCTGTATCA 1649
QY 1501 TGGAGGGCTTCTACCTTGTCTTTGATCGGGCCGAAACCAATTTGGCTTCTGTCGACGG 1560
Db 1650 TGGAGGGCTTCTACCTTGTCTTTGATCGGGCCGAAACCAATTTGGCTTCTGTCGACGG 1709

QY 1561 CTTGGCATGTGCACGATGAGTTTCAGGACGCGAGCGGTGGAAGGCCCTTTTGTCACTTGG 1620
Db 1710 CTTGGCATGTGCACGATGAGTTTCAGGACGCGAGCGGTGGAAGGCCCTTTTGTCACTTGG 1769
QY 1621 ACATGGAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG 1680
Db 1770 ACATGGAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCCCTCATGACCATAG 1829
QY 1681 CCTATGTATGCTGCCATCTGCCCTCTTTCATGCTGCCACTCTGCTCATGCTGTGTC 1740
Db 1830 CCTATGTATGCTGCCATCTGCCCTCTTTCATGCTGCCACTCTGCTCATGCTGTGTC 1889
QY 1741 AGTGGCGTGCCTCCGCTGCTCGCCAGCAGCATGATGACTTTTCTGCTGATGACATCTCCC 1800
Db 1890 AGTGGCGTGCCTCCGCTGCTCGCCAGCAGCATGATGACTTTTCTGCTGATGACATCTCCC 1949
QY 1801 TGCTGAAGTGGAGAGGCCCATGGCAGAAGATAGAGATTTCCCTGGAGCACACCTCCGTCG 1860
Db 1950 TGCTGAAGTGGAGAGGCCCATGGCAGAAGATAGAGATTTCCCTGGAGCACACCTCCGTCG 2009
QY 1861 GTTCACTTTGGTACAAAGTAGGAGACACAGATGCACCTGTGGCCAGAGCCTCAGGAC 1920
Db 2010 GTTCACTTTGGTACAAAGTAGGAGACACAGATGCACCTGTGGCCAGAGCCTCAGGAC 2069
QY 1921 CCTCCCCACCCACCAATGCCCTCTGCTTGTATGGAAGAAAGGCTGGCAAGGTGGT 1980
Db 2070 CCTCCCCACCCACCAATGCCCTCTGCTTGTATGGAAGAAAGGCTGGCAAGGTGGT 2129
QY 1981 TCCAGGGACTGTACCTGTAGGAACAGAAAGAGAAAGAGACACTCTGCTGGCGGA 2040
Db 2130 TCCAGGGACTGTACCTGTAGGAACAGAAAGAGAAAGAGACACTCTGCTGGCGGA 2189
QY 2041 ATACTCTTGGTCCACCTCAAAATTTAAGTCGGGAAATTTCTGCTTGAACCTTCAGCCCTG 2100
Db 2190 ATACTCTTGGTCCACCTCAAAATTTAAGTCGGGAAATTTCTGCTTGAACCTTCAGCCCTG 2249
QY 2101 AACCTTTGTCCACCAATTTCTTAAATTTCTCAACCCAAAGATTTCTCTTTTCTTACTTT 2160
Db 2250 AACCTTTGTCCACCAATTTCTTAAATTTCTCAACCCAAAGATTTCTCTTTTCTTACTTT 2309
QY 2161 CAGAGTACTTGGCATCACAGCAGGTTTACCTTGGCGTGTGTCCTTGGTGGTACCTTGGCAG 2220
Db 2310 CAGAGTACTTGGCATCACAGCAGGTTTACCTTGGCGTGTGTCCTTGGTGGTACCTTGGCAG 2369
QY 2221 AGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTC 2280
Db 2370 AGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTC 2429
QY 2281 CTATTTGCTTTAGAGACAGGACTGTATAAACAGCCCTTAACATTTGGTGCACAAAGATTGCCT 2340
Db 2430 CTATTTGCTTTAGAGACAGGACTGTATAAACAGCCCTTAACATTTGGTGCACAAAGATTGCCT 2489
QY 2341 CTTGAATT 2348
Db 2490 CTTGAATT 2497
RESULT 3
AB032975
LOCUS
DEFINITION Homo sapiens mRNA for KIAA1149 protein, partial cds.
ACCESSION AB032975
VERSION AB032975.2 GI:14133242
KEYWORDS
SOURCE Homo sapiens adult male brain cDNA to mRNA, clone_lib:pBluescriptII
SK plus clone:fg04087.
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Hirose, M., Nagase, T., Ishikawa, K., Kikuno, R., Nomura, N. and Ohara, O.

TITLE Characterization of cDNA clones selected by the Genemark analysis
JOURNAL DNA Res. 6 (5), 329-336 (1999)
MEDLINE 20039618
AUTHORS 2 (bases 1 to 5814)
TITLE Chara,O., Nagase,T. and Kikuno,R.
JOURNAL Direct Submission
Submitted (04-OCT-1999) Osamu Ohata, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp, Tel:+81-438-52-3913,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3914)
COMMENT On May 17, 2001 this sequence version replaced gi:6330044.
FEATURES Location/Qualifiers

Source
1..5814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fg04087"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
/dev_stage="adult"
/note="vector:pBluescriptII SK plus. This sequence is
replaced that of hg01289 cDNA as a representative cDNA
sequence for KIAA1149"
1..5814
/gene="KIAA1149"
351..1949
/gene="KIAA1149"
/codon_start=1
/note="Start codon is not identified."
/product="KIAA1149 protein"
/protein_id="BAA86463.2"
/db_xref="GI:14133243"
/translation="EATSTQTWGAQPGTDVGCPEPRGPPMPKPCPSGCCWA
RECCLPASTAGSPCAAAGAPLGLRPETDEEPEGRGFSFVEMVNLKGS
GGYVEMTPGPPQTLNILDVDTGSSNFAVGAAPLHRYQRLSSTYRLKGVY
VPTGGKRGELGLDLSVPHGNVNTVRANIAITSDKFFINGSNWEIGILAVAEI
APRDSLEPFDSLVHPLNLSQLQAGPPLNOSVLASVSGSMIIGDHSLY
TCSLWATPFRREYVIEIVRVEINGODLKMCKEYNDKSIYDSCGTTNLRPKVFE
AAVSKIKAASSTKEKFPDGFELGEQLVQWAGTTPWNIIPVLSILMLGVEVNSFRITI
LPQQLRPVEDVATSDDCYKFAISOSSTGVNAGVIMEGFYVVEDRARRKICFAVSA
CHVDFEFTAAVGGPFTVLDMDGYNIPQTDSTLMTIAYVMAATICALFMLPLCLMV
COMWRLCLRQGHDDFADDSLK"
BASE COUNT 1494 a 1510 c 1404 g 1406 t
ORIGIN

Query Match 99.3%; Score 2331.8; DB 9; Length 5814;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2344; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
QY 1 CCATGCCGCCCTCACAGCCCGCGGAGCCGAGCCCGCTGCCAGGCTGGCCCGCG 60
DB 138 CCATGCCGCCCTTCCAGCCCGCGGAGCCGCGCTGCCAGGCTGGCCCGCG 197
QY 61 CSCTGCCGATGTAGCGGCTCCGATCCAGCCTCTCCCTGCTCCCGTCTCGCGAT 120
DB 198 CCGTGCGATGTAGCGGCTCCGATCCAGCCTCTCCCTGCTCCCGTCTCGCGAT 257
QY 121 CTCCTCTGACCGCTTCCACAGCCCGGACCCGGGGGCTGCCCCAGGSCCTCGAGGCCCT 180
DB 258 CTCCTCTGACCGCTTCCACAGCCCGGACCCGGGGGCTGCCCCAGGSCCTCGAGGCCCT 317
QY 181 GCGCTCTGATGCCCCCAAGCTCCCTCTCTGAGAGCCAGCAGCACCAGACTTGG 240
DB 318 GCGCTCTGATGCCCCCAAGCTCCCTCTCTGAGAGCCAGCAGCACCAGACTTGG 377
QY 241 GGCAGGCGCCAGGAGCGACGTGGGCACTGCGAGCCAGAGGCGCCGAGGCGCGGCG 300
DB 378 GGCAGGCGCCAGGAGCGACGTGGGCACTGCGAGCCAGAGGCGCCGAGGCGCGGCG 437
QY 301 CCACCATGGCCCCAAGCCCTCCCTGGCTCTCTGTGGATGGGCGGCGAGTGTGCTG 360
|||||

DB 438 CCACCATGGCCCCAAGCCCTGCCCTGGCTCTGTGTGGATGGCGCGGAGTGTGCTG 497
QY 361 CCCACGGCACCCAGCAGCGCATCGGCTGCCCTGCCAGCGGCTT-GGGGGGGCGCCCC 419
DB 498 CCCACGGCACCCAGCAGCGCATCGGCTGCCCTGCCAGCGGCTGGGGGGCGCCCC 557
QY 420 CTGGGGTGGGCTGCCCTGCCCGGAGACCGAGAGAGCCGAGGAGCCCGCGGAGGGG 479
DB 558 CTGGGGTGGGCTGCCCTGCCCGGAGACCGAGAGAGCCCGAGGAGCCCGCGGAGGGG 617
QY 480 AGCTTTGTGAGATGTGTGACACCTGAGGGGCAAGTGTGGGGAGGCTACTAGTGGAG 539
DB 618 AGCTTTGTGAGATGTGTGACAACTGAGGGCAAGTGTGGGGAGGCTACTAGTGGAG 677
QY 540 ATGACGTGGGAGCCCGCCGAGAGCTCAACATCTCTGTGGATACAGCAGCAGTAAC 599
DB 678 ATGACGTGGGAGCCCGCCGAGAGCTCAACATCTCTGTGGATACAGCAGCAGTAAC 737
QY 600 TTTTGCAGTGGTGTGCTGCCCGCCACCCCTTCTTGCATCGCTACTACAGAGCAGCTGTCC 659
DB 738 TTTTGCAGTGGTGTGCTGCCCGCCACCCCTTCTTGCATCGCTACTACAGAGCAGCTGTCC 797
QY 660 AGCACATACCGGAGCTCCGAGAGGTGTGTATGTGCTTACACCCAGGCAAGTGGAA 719
DB 798 AGCACATACCGGAGCTCCGAGAGGTGTGTATGTGCTTACACCCAGGCAAGTGGAA 857
QY 720 GGGAGCTGGGAGCCGACCTGTTAAGCATCCCGATGCCCGCCCAACGTCACCTGTGCTGCC 779
DB 858 GGGAGCTGGGAGCCGACCTGTTAAGCATCCCGATGCCCGCCCAACGTCACCTGTGCTGCC 917
QY 780 AACATTGCTGCCATCACTGAATCAGCAAGTTCTTCAACAGCTCCCACTGGGAAGGC 839
DB 918 AACATTGCTGCCATCACTGAATCAGCAAGTTCTTCAACAGCTCCCACTGGGAAGGC 977
QY 840 ATCTGGGGTGGCTTATGTGTAGATTTGCAGGCTGACGACTCCCTGGAGCCTTCTTT 899
DB 978 ATCTGGGGTGGCTTATGTGTAGATTTGCAGGCTGACGACTCCCTGGAGCCTTCTTT 1037
QY 900 GACTCTCTGTAAAGCAGACCCAGTTCCCAACCTCTTCCCTGCGAGCTTTGTGGTGT 959
DB 1038 GACTCTCTGTAAAGCAGACCCAGTTCCCAACCTCTTCCCTGCGAGCTTTGTGGTGT 1097
QY 960 GGCTTCCCTTCAACCACTGTAAGTGTGCTGGCTTCTCGAGGAGGAGCATGATCATTTGA 1019
DB 1098 GGCTTCCCTTCAACCACTGTAAGTGTGCTGGCTTCTCGAGGAGGAGCATGATCATTTGA 1157
QY 1020 GGTATGACCACTCGCTGTACAGGAGTCTCTGGTATACACCATCCCGGCGGAGTGG 1079
DB 1158 GGTATGACCACTCGCTGTACAGGAGTCTCTGGTATACACCATCCCGGCGGAGTGG 1217
QY 1080 TATTATGAGTGTGATTTGCGGGTGGAGATCAATGAGAGGATCTGAAATGGAGTGC 1139
DB 1218 TATTATGAGTGTGATTTGCGGGTGGAGATCAATGAGAGGATCTGAAATGGAGTGC 1277
QY 1140 AAGAGTGTAACTATGACAGAGCATTTGTGACAGTGGCAGCACCACCACTTCTGTTTGGCC 1199
DB 1278 AAGAGTGTAACTATGACAGAGCATTTGTGACAGTGGCAGCACCACCACTTCTGTTTGGCC 1337
QY 1200 AGAAAGTGTGAGTGTGAGTCAATTCATCAAGGAGGCTCTCTCAAGGAGAGTTC 1259
DB 1338 AAGAAAGTGTGAGTGTGAGTCAATTCATCAAGGAGGCTCTCTCAAGGAGAGTTC 1397
QY 1260 CCTGATGTTTCTGGCTAGGAGAGCAGTGTGCTGGCAAGCAGCAGCACCCTTGG 1319
DB 1398 CCTGATGTTTCTGGCTAGGAGAGCAGTGTGCTGGCAAGCAGCAGCACCCTTGG 1457
QY 1320 AACATTTTCCAGTCACTCTACCTAATGGGTGAGGTTTACCAACCACTCTCTCCG 1379
DB 1458 AACATTTTCCAGTCACTCTACCTAATGGGTGAGGTTTACCAACCACTCTCTCCG 1517
QY 1380 ATCACCATCTTCCGAGCAATACCTGCGGCAAGTGTGGAAGATGTGGCCACATCCCAAGAC 1439
DB 1518 ATCACCATCTTCCGAGCAATACCTGCGGCAAGTGTGGAAGATGTGGCCACATCCCAAGAC 1577

Qy	1440	GACTGTTACAAGTTTGCCATCTCAGACTCATCCACGGGCACCTGTTATGGAGGCTGTTATC	1499
Db	1578	GACTGTTACAAGTTTGCCATCTCAGACTCATCCACGGGCACCTGTTATGGAGGCTGTTATC	1637
Qy	1500	ATGGAGGCTTCTACGTTGTCTTTTCATCGGGCCGAAAAACGAATGGCTTGTGCTGTGCAGC	1559
Db	1638	ATGGAGGCTTCTACGTTGTCTTTTCATCGGGCCGAAAAACGAATGGCTTGTGCTGTGCAGC	1697
Qy	1560	GCTTGCCATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAAGCCCTTTTGTACCTTG	1619
Db	1698	GCTTGCCATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAAGCCCTTTTGTACCTTG	1757
Qy	1620	GACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCTCATGACCATA	1679
Db	1758	GACATGGAAGACTGTGGCTACAACTTCCACAGACAGATGAGTCAACCTCATGACCATA	1817
Qy	1680	GCCATGTGTCATGGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCCTCATGGTGTGT	1739
Db	1818	GCCATGTGTCATGGCTGCCATCTGCGCCCTCTTTCATGCTGCCACTCTGCCTCATGGTGTGT	1877
Qy	1740	CAGTGGCCCTGCCTCCGCTGCCAGCAGCATGATGACTTTGCTGATGACATCTCC	1799
Db	1878	CAGTGGCCCTGCCTCCGCTGCCAGCAGCATGATGACTTTGCTGATGACATCTCC	1937
Qy	1800	CTGCTGAAGTCAGGAGGCCCATGCGGCAGAAGATAGAGATTCCCTCGACACACCTCCGT	1859
Db	1938	CTGCTGAAGTCAGGAGGCCCATGCGGCAGAAGATAGAGATTCCCTCGACACACCTCCGT	1997
Qy	1860	GTTTCACTTTTGGTCAAAATAGGAGACACAGATGGCACCTGTGGCCAGAGCACTTCAGGA	1919
Db	1998	GTTTCACTTTTGGTCAAAATAGGAGACACAGATGGCACCTGTGGCCAGAGCACTTCAGGA	2057
Qy	1920	CCCTCCCAACCCACCAATGCTCTGCCTTGATGGAGAAGAAAAGGCTGGCAAGTGGG	1979
Db	2058	CCCTCCCAACCCACCAATGCTCTGCCTTGATGGAGAAGAAAAGGCTGGCAAGTGGG	2117
Qy	1980	TTCCAGGACTGTACTCTGTAGAAACAGAAAAGAAAAGACACTCTGCTGCGCGG	2039
Db	2118	TTCCAGGACTGTACTCTGTAGAAACAGAAAAGAAAAGACACTCTGCTGCGCGG	2177
Qy	2040	AATACTTGTGGTCACTCAATTTAAGTCGGGAAATTCGTGCTGTTGAACTTCAGCCCT	2099
Db	2178	AATACTTGTGGTCACTCAATTTAAGTCGGGAAATTCGTGCTGTTGAACTTCAGCCCT	2237
Qy	2100	GAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGATCTCTCTTTTCTTAGTT	2159
Db	2238	GAACCTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGATCTCTCTTTTCTTAGTT	2297
Qy	2160	TCAGAACTACTGGCATCACACGAGGTTACCTTGGCGTGTGCCCTGTGGTAGCCCTGGCA	2219
Db	2298	TCAGAACTACTGGCATCACACGAGGTTACCTTGGCGTGTGCCCTGTGGTAGCCCTGGCA	2357
Qy	2220	GAGAAGACCAAGCTTGTTCCTTCGTCGGCCAAAGTCAGTAGGAGAGATGCACAGTTT	2279
Db	2358	GAGAAGACCAAGCTTGTTCCTTCGTCGGCCAAAGTCAGTAGGAGAGATGCACAGTTT	2417
Qy	2280	GCTATTTCCTTTAGACAGGACTGTATAAACAAGCCCTAACTTGGTGCACAAAGATTGCC	2339
Db	2418	GCTATTTCCTTTAGACAGGACTGTATAAACAAGCCCTAACTTGGTGCACAAAGATTGCC	2477
Qy	2340	TCCTGAATT 2348	
Db	2478	TCCTGAATT 2486	

RESULT 4					
AX364933					
LOCUS					
DEFINITION	AX364933	5757 bp	DNA	linear	PAT 15-FEB-2002
ACCESSION	AX364933	Sequence 84	from Patent	WO0206315.	
VERSION	AX364933.1	GI:18696823			
KEYWORDS	.				

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 AUTHORS Mintz, L., Freilich, S. and Bernstein, J.
 TITLE Novel nucleic acid and amino acid sequences
 JOURNAL Patent: WO 0206315-A 84 24-JAN-2002;
 Compugen Ltd. (IL)
 FEATURES
 source Location/Qualifiers
 1..5757
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1485 a 1490 c 1392 g 1388 t 2 others
 ORIGIN
 Query Match 92.98; Score 2181.2; DB 6; Length 5757;
 Best Local Similarity 96.68; Pred. No. 0;
 Matches 2268; Conservative 1; Mismatches 4; Indels 75; Gaps 1;
 QY 1 CCATGCGGGCCCTCACAGCCCCCGGAGCGCCGCTGCCAGAGCTGGCCGCCG 60
 DB 150 CCATGCGGGCCCTCCACAGCCCCCGGAGCGCCGCTGCCAGAGCTGGCCGCCG 209
 QY 61 CSGTGCCGATGTAGCGGGCTCCGGATCCAGAGCTCTCCCTGCTGCCGTGCTGCGGAT 120
 DB 210 CSGTGCCGATGTAGCGGGCTCCGGATCCAGAGCTCTCCCTGCTGCCGTGCTGCGGAT 269
 QY 121 CTCCTCTGACCGTCTCCACAGCCGAGCCGGGGCTGGCCCCAGGGCCCTGCAGGCCCT 180
 DB 270 CTCCTCTGACCGTCTCCACAGCCGAGCCGGGGCTGGCCCCAGGGCCCTGCAGGCCCT 329
 QY 181 GGGCTCTGATGCCCCCAAGCTCCCTCTCTGAGAAGCCACGACACCCAGACTGG 240
 DB 330 GGGCTCTGATGCCCCCAAGTCCCTCTCTGAGAAGCCACGACACCCAGACTGG 389
 QY 241 GGGCAGCGCCAGGACGAGCTGGCCAGTGGCAGCCAGAGGCCCGGAAGCGCGGGC 300
 DB 390 GGGCAGCGCCAGGACGAGCTGGCCAGTGGCAGCCAGAGGCCCGGAAGCGCGGGC 449
 QY 301 CCACCATGGCCCAAGCCCTGCCCTGGCTCTGCTGAGTGGCGGGGAGTCTGCCTG 360
 DB 450 CCACCATGGCCCAAGCCCTGCCCTGGCTCTGCTGAGTGGCGGGGAGTCTGCCTG 509
 QY 361 CCCACGSCACCCAGCACGGCATCGGCTGCCCTTCGCGACGCGCTTGGGGGGCGCCCC 420
 DB 510 CCCACGSCACCCAGCACGGCATCGGCTGCCCTTCGCGACGCGCTTGGGGGGCGCCCC 569
 QY 421 TGGGGCTGCGGGCTGCCCCGGGAGACGACGAGAGCCCGAGGAGCCCGCGGAGGGCA 480
 DB 570 TGGGGCTGCGGGCTGCCCCGGGAGACGACGAGAGAGCCCGAGGAGCCCGCGGAGGGCA 629
 QY 481 GCCTTGTGGAGATGGTGACAACTGAGGGCAAGTGGGGGAGGCTACTACGTGAGA 540
 DB 630 GCCTTGTGGAGATGGTGACAACTGAGGGCAAGTGGGGGAGGCTACTACGTGAGA 689
 QY 541 TGACCTGTGGCAGCCCCCGCGAGACGCTCAACATCCTTGGTGATACAGCAGCAGTAAC 600
 DB 690 TGACCTGTGGCAGCCCCCGCGAGACGCTCAACATCCTTGGTGATACAGCAGCAGTAAC 749
 QY 601 TTGCACTGGGTGTGCCCCCGACCCCTTCCTGCATCGCTACTACAGAGGCAGCTGCCA 660
 DB 750 TTGCACTGGGTGTGCCCCCGACCCCTTCCTGCATCGCTACTACAGAGGCAGCTGCCA 809
 QY 661 GCACATACCGGGACCTCCGGAGGGTGTATGTGCCCTACACCCAGGCGAAGTGGGAAG 720
 DB 810 GCACATACCGGGACCTCCGGAGGGTGTATGTGCCCTACACCCAGGCGAAGTGGGAAG 869
 QY 721 GGGAGCTGGGACCGGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACCTGCGTGCCA 780
 DB 870 GGGAGCTGGGACCGGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACCTGCGTGCCA 929
 QY 781 ACATTGCTGCATCAGTGAATCAGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCA 840

|||||
930 ACATTGCTGCTCATTCACTGAATCAGAAAGTTCTTCAATCAACGGCTCCAACCTGGAAGGCA 989
QY TCCTGGGGCTGGCCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTCTTTG 900
Db TCCTGGGGCTGGCCTATGCTGAGATTGCCAG----- 1020
QY ACTCTCTGGTAAAGCAGACACCCACAGTTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGGCTG 960
Db 1021 -----GCTTTGTGGTGGCTG 1034
QY GCTTCCCTCCACACAGTCTGAAGTGTGCGCTCTGTGCGAGGAGGAGCATGATCATTTGGAG 1020
Db 1035 GCTTCCCTCCACACAGTCTGAAGTGTGCGCTCTGTGCGAGGAGGAGCATGATCATTTGGAG 1094
QY 1021 GTATCGACCACTCGCTGTACACAGGAGTCTCTGGTATACACCATCCGCGGGAGTGGT 1080
Db 1095 GTATCGACCACTCGCTGTACACAGGAGTCTCTGGTATACACCATCCGCGGGAGTGGT 1154
QY 1081 ATTTATGAGGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCA 1140
Db 1155 ATTTATGAGGTGATCATTTGTGCGGTGGAGATCAATGGACAGGATCTGAAATGGACTGCA 1214
QY 1141 AGGAGTACAACATATGACAAGAGCATTTGTGGACAGTGGCACCAACACCTTTGCTTTGCCCA 1200
Db 1215 AGGAGTACAACATATGACAAGAGCATTTGTGGACAGTGGCACCAACACCTTTGCTTTGCCCA 1274
QY 1201 AGAAGTGTTTGAAGCTGCAAGTCAAAATCCATCAAGCAGCTCTCCACCGGAGAGTTCC 1260
Db 1275 AGAAGTGTTTGAAGCTGCAAGTCAAAATCCATCAAGCAGCTCTCCACCGGAGAGTTCC 1334
QY 1261 CTGATGTTTCTGGCTAGGAGCAGCTGCTGTGCTGGCAAGCAGGACCAACCCCTTTGGA 1320
Db 1335 CTGATGTTTCTGGCTAGGAGCAGCTGCTGTGCTGGCAAGCAGGACCAACCCCTTTGGA 1394
QY 1321 ACATTTTCCAGTCACTCACTCACTAATGGGTGAGTTTACAAACCACTGCTTCTCGCA 1380
Db 1395 ACATTTTCCAGTCACTCACTCACTAATGGGTGAGTTTACAAACCACTGCTTCTCGCA 1454
QY 1381 TCACCATCTCTCCGACGAATACCTCGCGCAGTGAAGATGTGGCCACCTGCCAAGACG 1440
Db 1455 TCACCATCTCTCCGACGAATACCTCGCGCAGTGAAGATGTGGCCACCTGCCAAGACG 1514
QY 1441 ACTGTTACAAGTTTGGCATCTCAGTCACTCCACGGCAGCTGTATGGGAGCTGTTATCA 1500
Db 1515 ACTGTTACAAGTTTGGCATCTCAGTCACTCCACGGCAGCTGTATGGGAGCTGTTATCA 1574
QY 1501 TGGAGGCTTCTACGTTGCTTTGATCGGCGCCGAAACGAATTTGCTTGTCTGCTCAGCG 1560
Db 1575 TGGAGGCTTCTACGTTGCTTTGATCGGCGCCGAAACGAATTTGCTTGTCTGCTCAGCG 1634
QY 1561 CTTGCCATGTGCACGATGAGTTTCAGGACGCGCAGGTTGGAAGGCCCTTTTGTCACTTGG 1620
Db 1635 CTTGCCATGTGCACGATGAGTTTCAGGACGCGCAGGTTGGAAGGCCCTTTTGTCACTTGG 1694
QY 1621 ACATGAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCTCATGACCATAG 1680
Db 1695 ACATGAAGACTGTGGCTACAACATTTCCACAGACAGATGAGTCAACCTCATGACCATAG 1754
QY 1681 CCTATGTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTCTGCTCATGCTGTGTC 1740
Db 1755 CCTATGTCATGGCTGCCATCTGGCCCTCTTTCATGCTGCCACTCTGCTCATGCTGTGTC 1814
QY 1741 AGTGGCGCTCCCTCCGCTGCTGCGCAGCAGCATGATGATCTTGTGATGACATCTCCC 1800
Db 1815 AGTGGCGCTCCCTCCGCTGCTGCGCAGCAGCATGATGATCTTGTGATGACATCTCCC 1874
QY 1801 TGCTGAAGTGAGGAGGCCCTTGGCAGAGATAGAGATTCCTCCCTGGACCAACCTCCGCTG 1860
Db 1875 TGCTGAAGTGAGGAGGCCCTTGGCAGAGATAGAGATTCCTCCCTGGACCAACCTCCGCTG 1934
QY 1861 GTTTCATCTTGGTCACAAGTAGGAGACACAGATGGCACCTGCTGGCCAGAGCACTCAGGAC 1920
|||||

Db 1935 GTTCACCTTTGGTCACAAAGTAGGAGACACAGATGGCAGCTGTGGCCAGACACCTCAGGAC 1994
QY 1921 CTCTCCACCCACCAAAATGCTCTGCTTGTATGGAGAGAAAAGGCTGGCAAGGTGGT 1980
Db 1995 CTCTCCACCCACCAAAATGCTCTGCTTGTATGGAGAGAAAAGGCTGGCAAGGTGGT 2054
QY 1981 TCCAGGAGCTGTACCTGTAGGAACAGAAAGAGAGAAAGCACTCTGCTGGCGGGA 2040
Db 2055 TCCAGGAGCTGTACCTGTAGGAACAGAAAGAGAGAAAGCACTCTGCTGGCGGGA 2114
QY 2041 ATACTCTTGTGCTCACTCAAAATTTAAGTCGGGAAATTTCTGCTGTTGAACTTCAGCCCTG 2100
Db 2115 ATACTCTTGTGCTCACTCAAAATTTAAGTCGGGAAATTTCTGCTGTTGAACTTCAGCCCTG 2174
QY 2101 AACCTTTGTCCACCATCTTAAATTTCTCAACCCAAAGTATTCTTTCTTTCTTAGTTT 2160
Db 2175 AACCTTTGTCCACCATCTTAAATTTCTCAACCCAAAGTATTCTTTCTTTCTTAGTTT 2234
QY 2161 CAGAAGTACTGGCATCACACGAGTTTACCTTGGCGTGTGTCCCTGTGGTACCCCTGGCAG 2220
Db 2235 CAGAAGTACTGGCATCACACGAGTTTACCTTGGCGTGTGTCCCTGTGGTACCCCTGGCAG 2294
QY 2221 AGAAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG 2280
Db 2295 AGAAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTG 2354
QY 2281 CTATTTCTTTAGAGACAGGAGCTGTATAAACAAGCCTTAACATTTGGTGCAGAAAGATTGCT 2340
Db 2355 CTATTTCTTTAGAGACAGGAGCTGTATAAACAAGCCTTAACATTTGGTGCAGAAAGATTGCT 2414
QY 2341 CTTGAATT 2348
Db 2415 CTTGAATT 2422

RESULT 5
AF200343 Homo sapiens chromosome 11 aspartyl protease 2 mRNA, complete cds.
LOCUS AF200343
DEFINITION AF200343
ACCESSION AF200343
VERSION AF200343.1 GI:6561813
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 2070)
Yan, R., Blenkowski, M.J., Shuck, M.E., Miao, H., Tory, M.C.,
Pauley, A.M., Brashier, J.R., Stratman, N.C., Mathews, W.R., Buhl, A.E.,
Carter, D.B., Tomasselli, A.G., Parodi, L.A., Heinrikson, R.L. and
Gurney, M.E.

TITLE
Membrane-anchored aspartyl protease with Alzheimer's disease
beta-secretase activity
Nature 402 (6761), 533-537 (1999)
JOURNAL
MEDLINE 20057170
PUBMED 10591213
REFERENCE
2 (bases 1 to 2070)
Blenkowski, M.J., Shuck, M.E., Slightom, J.L. and Drong, R.F.
AUTHORS
Direct Submission
TITLE
JOURNAL
Submitted (29-Oct-1999) Genomics Research, Pharmac&Upjohn, 301
Henrietta, Kalamazoo, MI 49007, USA
FEATURES
Location/Qualifiers
1..2070
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q23.3-24.1"
1..1506
/codon_start=1
/product="aspartyl protease 2"
/protein_id="AAFI709.1"
/db_xref="GI:6561814"
/translation="MAQALPWLIIWMGAGVLPAGHTQHGIRLPLRSGLGAPLGRLP

QY 2286 TGCTTTAGACAGGACTGTATATAACAGCCCTAACATTGGTGCACAGATTGGCTCTTGA 2345
|||||
Db 1981 TGCTTTAGACAGGACTGTATATAACAGCCCTAACATTGGTGCACAGATTGGCTCTTGA 2040

QY 2346 ATT 2348
|||
Db 2041 ATT 2043

RESULT 6
AX105385 LOCUS AX105385 2070 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123533.
ACCESSION AX105385
VERSION AX105385.1 GI:13921511
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2070)
AUTHORS Gurney,M. and Bienkowski,M.J.
TITLE Alzheimer's disease secretase, app substrates therefor, and uses therefor
JOURNAL Patent: WO 0123533-A 3 05-APR-2001;
Pharmacia & Upjohn Company (US)
FEATURES
source Location/Qualifiers
1..2070
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 476 a 583 c 562 g 449 t
ORIGIN

Query Match 86.9%; Score 2039.8; DB 6; Length 2070;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 306 ATGGCCCAAGCCCTGCCCTGGCTCTGCTGTGATGGCGCGGAGTGTGCTGCCAC 365
Db 1 ATGGCCCAAGCCCTGCCCTGGCTCTGCTGTGATGGCGCGGAGTGTGCTGCCAC 60

QY 366 GGCACCCAGCACGGCATCGCGCTGCCCTGCGCAGCGGCGTGGGGCGCCCTGGGG 425
Db 61 GGCACCCAGCACGGCATCGCGCTGCCCTGCGCAGCGGCGTGGGGCGCCCTGGGG 120

QY 426 CTCGGCTGCCCGGAGACCCAGACGAGACCCGAGGAGCCCGCGGGAGGCGACGTTT 485
Db 121 CTCGGCTGCCCGGAGACCCAGACGAGACCCGAGGAGCCCGCGGGAGGCGACGTTT 180

QY 486 GTGGAGATGGTGACAACTGAGGGCAAGTTCGGGCGAGGCTACTACGTGGAGATGACC 545
Db 181 GTGGAGATGGTGACAACTGAGGGCAAGTTCGGGCGAGGCTACTACGTGGAGATGACC 240

QY 546 GTGGGAGCCCCCGGAGAGCGCTCAACATCTCTGGTGGATACAGCGACGACTTGTGCA 605
Db 241 GTGGGAGCCCCCGGAGAGCGCTCAACATCTCTGGTGGATACAGCGACGACTTGTGCA 300

QY 606 GTGGGTGCTGCCCGCCCGCCCGCTTCTGCTGATCGCTACTACAGAGGAGCTGTCCAGCACA 665
Db 301 GTGGGTGCTGCCCGCCCGCCCGCTTCTGCTGATCGCTACTACAGAGGAGCTGTCCAGCACA 360

QY 666 TACCGGACCTCCGGAGGGTGTGTGTGCCCTTACACCCAGGGCAAGTGGGAAGGGGAG 725
Db 361 TACCGGACCTCCGGAGGGTGTGTGTGCCCTTACACCCAGGGCAAGTGGGAAGGGGAG 420

QY 726 CTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCCAACCTCACTGTGCGTGGCAACATT 785
Db 421 CTGGGACCGACCTGGTAAGCATCCCCCATGGCCCCCAACCTCACTGTGCGTGGCAACATT 480

QY 786 GCTGGCATCACTGAATCAGACAGATTCTTTCATCAAGGCTTCCACCTGGGAGGAGCATCCTG 845
Db 481 GCTGGCATCACTGAATCAGACAGATTCTTTCATCAAGGCTTCCACCTGGGAGGAGCATCCTG 540

QY 846 GGCTGGCCTATGCTGAGATTGCCAGGCTTGACGACTCCCTGGAGCCTTTCTTTGACTCT 905
|||||
Db 541 GGCTGGCCTATGCTGAGATTGCCAGGCTTGACGACTCCCTGGAGCCTTTCTTTGACTCT 600

QY 906 CTGGTAAAGCAGACCCACAGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGTGTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACAGTTCCCAACCTCTTCTCCCTGCAGCCTTTGTGTGTGGCTTC 660

QY 966 CCCCTCAACAGTCTGAAGTGTGGCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC 1025
|||||
Db 661 CCCCTCAACAGTCTGAAGTGTGGCTCTGTGCGAGGGAGCATGATCATTTGGAGGTATC 720

QY 1026 GACCACTCGCTGTACACAGGAGTCTCTGTGTATACACCCATCCGGCGGAGTGTATTAT 1085
Db 721 GACCACTCGCTGTACACAGGAGTCTCTGTGTATACACCCATCCGGCGGAGTGTATTAT 780

QY 1086 GAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGATCTGAAATGGACTGCAAGGAG 1145
|||||
Db 781 GAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGATCTGAAATGGACTGCAAGGAG 840

QY 1146 TACAACATATGACAGAGCATTTGGAGAGTGGCAGTGGCACCACCACTTCTGTTTGGCCCAAGAAA 1205
Db 841 TACAACATATGACAGAGCATTTGGAGAGTGGCAGTGGCACCACCACTTCTGTTTGGCCCAAGAAA 900

QY 1206 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGAGGCTCTCTCCACGGAGAAAGTTCCCTGAT 1265
Db 901 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGGAGGCTCTCTCCACGGAGAAAGTTCCCTGAT 960

QY 1266 GGTTCCTGGCTAGGAGAGCAGTGGTGTGTGGCAGCAGGACCAACCCCTTGGAAACATT 1325
Db 961 GGTTCCTGGCTAGGAGAGCAGTGGTGTGTGGCAGCAGGACCAACCCCTTGGAAACATT 1020

QY 1326 TTCCCACTCATCTCACTCTACCTTAATGGGTGAGGTTACCAACCACTCTTCCCGCATCACC 1385
Db 1021 TTCCCACTCATCTCACTCTACCTTAATGGGTGAGGTTACCAACCACTCTTCCCGCATCACC 1080

QY 1386 ATCTCTCCGAGCAATATACCTTGGCCAGTGGAGATGTGGCCACGTCCTCCCAAGAGCACTGT 1445
Db 1081 ATCTCTCCGAGCAATATACCTTGGCCAGTGGAGATGTGGCCACGTCCTCCCAAGAGCACTGT 1140

QY 1446 TACAAGTTTCCCATCTCACAGTCAATCCACGGGCACTGTATGGAGAGCTGTATATCATGGAG 1505
Db 1141 TACAAGTTTCCCATCTCACAGTCAATCCACGGGCACTGTATGGAGAGCTGTATATCATGGAG 1200

QY 1506 GCTTCTACGTTGTCTTTGATCGGGCCGAAAACGAATTTGGCTTGTGTCAGGCGCTTGC 1565
Db 1201 GCTTCTACGTTGTCTTTGATCGGGCCGAAAACGAATTTGGCTTGTGTCAGGCGCTTGC 1260

QY 1566 CATGTGCACGATGAGTTTCAGGAGCGGAGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1625
Db 1261 CATGTGCACGATGAGTTTCAGGAGCGGAGCGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1320

QY 1626 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
Db 1321 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380

QY 1686 GTCATGGCTGCCATTCGCGCCCTTTCATGCTGCCACTTGCCTCATGTTGTGTGTCAGTGG 1745
Db 1381 GTCATGGCTGCCATTCGCGCCCTTTCATGCTGCCACTTGCCTCATGTTGTGTGTCAGTGG 1440

QY 1746 CGTGTCCCTCGCTGCCCTTGCAGGACGATGATGACATTTGCTGATGACATCTCCCTGCTG 1805
Db 1441 CGTGTCCCTCGCTGCCCTTGCAGGACGATGATGACATTTGCTGATGACATCTCCCTGCTG 1500

QY 1806 AAGTGAAGAGGCCCATGGGAGAGATAGAGATTCCTCCCTGGACACACCTCCCGTGGTTTCA 1865
Db 1501 AAGTGAAGAGGCCCATGGGAGAGATAGAGATTCCTCCCTGGACACACCTCCCGTGGTTTCA 1560

QY 1866 CTTTGGTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCACTCAGGACCCCTCC 1925
Db 1561 CTTTGGTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCACTCAGGACCCCTCC 1620

Db	1261																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BASE COUNT	598 a	673 c	675 g	579 t	16 others
ORIGIN					
Query Match	86.2%; Score 2024; DB 6; Length 2541;				
Best Local Similarity	99.7%; Pred. No. 0;				
Matches 2038; Conservative	0; Mismatches 5; Indels 1; Gaps 1;				
QY	306	ATGCCCCAAGCCCTGCCTGCCTGCTGTGGATGGCGCGGAGTGTGCTGCTGCCAC	365		
Db	1	ATGCCCCAAGCCCTGCCTGCCTGCTGTGGATGGCGCGGAGTGTGCTGCTGCCAC	60		
QY	366	GGCACCCAGCACGGCATCCGGCTGCCCTGGCAGCGCCCTGGGGGCGCCCCCTGGG	425		
Db	61	GGCACCCAGCACGGCATCCGGCTGCCCTGGCAGCGCCCTGGGGGCGCCCCCTGGG	120		
QY	426	CTGGGCTGCCCGGGAGACGCGAGAGAGCCCGAGAGCCCGCCGAGGAGGAGCTTT	485		
Db	121	CTGGGCTGCCCGGGAGACGCGAGAGAGCCCGAGAGCCCGCCGAGGAGGAGCTTT	180		
QY	486	GTGGAGATGTGGACAACTTGAGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC	545		
Db	181	GTGGAGATGTGGACAACTTGAGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC	240		
QY	546	GTGGGACGCCCCCGCAGACGCTCAACATCCTGTGTGATACAGGCACAGTAACCTTTCGA	605		
Db	241	GTGGGACGCCCCCGCAGACGCTCAACATCCTGTGTGATACAGGCACAGTAACCTTTCGA	300		
QY	606	GTGGGTCTGCCCGCCCGCCCTTCTGCATCGCTACTACCAGAGGAGCTGTCCAGCACA	665		
Db	301	GTGGGTCTGCCCGCCCGCCCTTCTGCATCGCTACTACCAGAGGAGCTGTCCAGCACA	360		
QY	666	TACCGGACCTCCGGAAGGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG	725		
Db	361	TACCGGACCTCCGGAAGGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGGAG	420		
QY	726	CTGGGACCCGACCTGTGAAGCATCCCCATGGCCCCAACGCTCACTGTGCGTGCCAACATT	785		
Db	421	CTGGGACCCGACCTGTGAAGCATCCCCATGGCCCCAACGCTCACTGTGCGTGCCAACATT	480		
QY	786	GCTGCCATCACTGAATCAGACAAGTTCTTATCAACGGCTCCAACTGGGAAGGATCTCTG	845		
Db	481	GCTGCCATCACTGAATCAGACAAGTTCTTATCAACGGCTCCAACTGGGAAGGATCTCTG	540		
QY	846	GGGCTGCGCTATGTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT	905		
Db	541	GGGCTGCGCTATGTGAGATTGCCAGGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT	600		
QY	906	CTGCTAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGACGCTTTGTGTGCTGGCTTC	965		
Db	601	CTGCTAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGACGCTTTGTGTGCTGGCTTC	660		
QY	966	CCCTCAACGAGTCTGAAGTCTGCGCTCTCTCGGAGGAGCATGATCATTTGGAGGTATC	1025		
Db	661	CCCTCAACGAGTCTGAAGTCTGCGCTCTCTCGGAGGAGCATGATCATTTGGAGGTATC	720		
QY	1026	GACCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGCGGAGTGTATTAT	1085		
Db	721	GACCACTCGCTGTACACAGGAGTCTCTGGTATACACCCATCCCGCGGAGTGTATTAT	780		
QY	1086	GAGGTGATCATTTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAAATGGACTGCAAGAG	1145		
Db	781	GAGGTGATCATTTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAAATGGACTGCAAGAG	840		
QY	1146	TACAACTATGACAGAGCATTTGTGGACAGTGGCACCAACCTTCGTTTCCCCAAGAAA	1205		
Db	841	TACAACTATGACAGAGCATTTGTGGACAGTGGCACCAACCTTCGTTTCCCCAAGAAA	900		
QY	1206	GTGTTTGAAGCTGAGTCAAAATCCATCAAGGACGCTCTCCACGGAGAACTTCCCTGAT	1265		
Db	901	GTGTTTGAAGCTGAGTCAAAATCCATCAAGGACGCTCTCTCCACGGAGAACTTCCCTGAT	960		
QY	1266	GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGACCAACCCCTTGGAAACATT	1325		

Db 961 GGTTCGTGGCTAGAGAGCAGCTGGTGTGCTGGCAGAGGACCCACCTTGGAAACATT 1020
QY 1326 TTCCAGCTCATCTCACTTACCTAATGGGTGAGGTTACCAACAGTCTTCCGCATCACC 1385
Db 1021 TTCCAGCTCATCTCACTTACCTAATGGGTGAGGTTACCAACAGTCTTCCGCATCACC 1080
QY 1386 ATCTTCCGGAGCAATACCTCGGGCCAGTGGGAAGATGTGGCCACAGTCCCAAGACGACTGT 1445
Db 1081 ATCTTCCGGAGCAATACCTCGGGCCAGTGGGAAGATGTGGCCACAGTCCCAAGACGACTGT 1140
QY 1446 TACAAGTTTGGCATCTCACAGTCTCCACGGCAGCTGTTATGGAGCTGTTATCATGGAG 1505
Db 1141 TACAAGTTTGGCATCTCACAGTCTCCACGGCAGCTGTTATGGAGCTGTTATCATGGAG 1200
QY 1506 GGCTTTCACGTTGCTTTGATCGGGCCCGAAAACGAATTTGCTGTGCAGCGTTGC 1565
Db 1201 GGCTTTCACGTTGCTTTGATCGGGCCCGAAAACGAATTTGCTGTGCAGCGTTGC 1260
QY 1566 CATGTGCACGATGATTCAGGACGCGACGGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1625
Db 1261 CATGTGCACGATGATTCAGGACGCGACGGGTGGAAGGCCCTTTTGTACCTTTGGACATG 1320
QY 1626 GAAGACTGTGCTACAAACATTCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1685
Db 1321 GAAGACTGTGCTACAAACATTCACAGACAGATGAGTCAACCTCATGACCATAGCCTAT 1380
QY 1686 GTCATGGCTGCCATCTGCGCCCTCTTTCATGTGCCACTCTGCTCATGTGTGTCAAGTGG 1745
Db 1381 GTCATGGCTGCCATCTGCGCCCTCTTTCATGTGCCACTCTGCTCATGTGTGTCAAGTGG 1440
QY 1746 CGCTGCCTCGCTCGCTGCGCCAGCAGCATGATGATTTGCTGTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGCCTCGCTCGCTGCGCCAGCAGCATGATGATTTGCTGTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGGAGGCGCCATGGGCAGAGATAGAGATTCCCTT -GGACCAACCTCCCTGCTGTTTC 1864
Db 1501 AAGTGGAGGCGCCATGGGAAAGATAGAGATTCCCTTGGGACCACTCCCTGCTGTTTC 1560
QY 1865 ACTTTGGTCAACAAGTAGGAGACACAGATGGCACCTGTGSCCAGAGACACCTTCAGGACCCCTC 1924
Db 1561 ACTTTGGTCAACAAGTAGGAGACACAGATGGCACCTGTGSCCAGAGACACCTTCAGGACCCCTC 1620
QY 1925 CCACACCACCAATGCTCTGCTTGTAGGAGAGGAAAGGCTGGCAAGTGGGTGCTTCCA 1984
Db 1621 CCACACCACCAATGCTCTGCTTGTAGGAGAGGAAAGGCTGGCAAGTGGGTGCTTCCA 1680
QY 1985 GGGACTCTACTGTAGGAAACAGAAACAGAGAAAGAGCACTCTGCTGGCGGGGAATAC 2044
Db 1681 GGGACTCTACTGTAGGAAACAGAAAGAGAAAGAGCACTCTGCTGGCGGGGAATAC 1740
QY 2045 TCTTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTGCTTGAACCTTCAGCCCTGAACC 2104
Db 1741 TCTTGGTCACTCAAAATTTAAGTCGGGAAATTCGCTGCTTGAACCTTCAGCCCTGAACC 1800
QY 2105 TTTGTCCACCAATTCCTTTAAATTCCTCAACCCAAAGTATCTCTTTCTTAGTTTTCAGA 2164
Db 1801 TTTGTCCACCAATTCCTTTAAATTCCTCAACCCAAAGTATCTCTTTCTTAGTTTTCAGA 1860
QY 2165 AGTACTGGCATCACAGCAGGTTTACCTTTGGCGTGTGTCCTGTGGTACCTGGCAGAGAA 2224
Db 1861 AGTACTGGCATCACAGCAGGTTTACCTTTGGCGTGTGTCCTGTGGTACCTGGCAGAGAA 1920
QY 2225 GAGACCAAGCTTGTTCCTCTGCTGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGTCTAT 2284
Db 1921 GAGACCAAGCTTGTTCCTCTGCTGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGTCTAT 1980
QY 2285 TTGCTTTAGACAGAGGACTGTATAAACAAGCCTAACATTGGTGCAAGATTCCTCTCTTG 2344
Db 1981 TTGCTTTAGACAGAGGACTGTATAAACAAGCCTAACATTGGTGCAAGATTCCTCTCTTG 2040
QY 2345 AATT 2348
Db 2041 AATT 2044

RESULT 9
AX002655
LOCUS AX002655 2541 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 1 from Patent EP0855444.
ACCESSION AX002655
VERSION AX002655.1 GI:7242133
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 2541)
AUTHORS Murphy,K. and Chapman,C.G.
TITLE Aspartic proteinase 2 (ASP2)
JOURNAL Patent: EP 0855444-A 1 29-JUL-1998;
SMITHKLINE BEECHAM PLC (GB); SMITHKLINE BEECHAM CORP (US)
FEATURES
source
1..2541
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 598 a 673 c 675 g 579 t 16 others
ORIGIN
Query Match 86.2%; Score 2024; DB 6; Length 2541;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2038; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
Qy 306 ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGGGAGTCTGCCTGCCAC 365
Db 1 ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGGGAGTCTGCCTGCCAC 60
Qy 366 GGCACCCACGACGGCATCCGGCTGCCCTGCGACGGCCCTGGGGGGCGCCCTCGGG 425
Db 61 GGCACCCACGACGGCATCCGGCTGCCCTGCGACGGCCCTGGGGGGCGCCCTCGGG 120
Qy 426 CTGGCGCTGCCCGGGAGACCGAGAGAGCCCGAGAGCCCGGGGGAGGGGAGCTTT 485
Db 121 CTGGCGCTGCCCGGGAGACCGAGAGAGCCCGAGAGCCCGGGGGAGGGGAGCTTT 180
Qy 486 GTGGAGATGTGGACAACCTGAGGGCAAGTCGGGGAGGGCTACTAGCTGGAGATGACC 545
Db 181 GTGGAGATGTGGACAACCTGAGGGCAAGTCGGGGAGGGCTACTAGCTGGAGATGACC 240
Qy 546 GTGGGAGCCCGCCCGCAGACGCTCAACATCCTGTGTGATACAGGCAGCAGTAACCTTGC 605
Db 241 GTGGGAGCCCGCCCGCAGACGCTCAACATCCTGTGTGATACAGGCAGCAGTAACCTTGC 300
Qy 606 GTGGGTGTGCGCCCGCCACCCCTTCTGTCATGCTGTACTACAGAGCAGCTGTCCAGACA 665
Db 301 GTGGGTGTGCGCCCGCCACCCCTTCTGTCATGCTGTACTACAGAGCAGCTGTCCAGACA 360
Qy 666 TACCGGGAGCTCCGGAAGGCTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAG 725
Db 361 TACCGGGAGCTCCGGAAGGCTGTATGTAGCCCTACACCCAGGCAAGTGGGAAGGGAG 420
Qy 726 CTGGGCACCGACCTGTGTAGCATCCCGCATGGCCCCAACAGTCACCTGTGCTGCCAACAT 785
Db 421 CTGGGCACCGACCTGTGTAGCATCCCGCATGGCCCCAACAGTCACCTGTGCTGCCAACAT 480
Qy 786 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCTTG 845
Db 481 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCTTG 540
Qy 846 GGGCTGGCCCTATGCTCAGATTGCCAGGCTTACCACCTCCCTGGAGCCCTTCTTTGACTCT 905
Db 541 GGGCTGGCCCTATGCTCAGATTGCCAGGCTTACCACCTCCCTGGAGCCCTTCTTTGACTCT 600
Qy 906 CTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCTCCCTGACAGCTTTGTGGTCTGGCTTC 965
Db 601 CTGGTAAAGCAGACCCACGCTTCCCAACCTCTTCTCCCTGACAGCTTTGTGGTCTGGCTTC 660
Qy 966 CCCCTCAACCAAGTCTGAAGTCTGGCCCTCTGTCCGAGGGAGCATGATCATTTGGAGTATC 1025

||||| 661 CCCTCAACAGCTCAAGTGGCTGGCTCTGTGGAGGGAGCATGATCATTTGGAGGTATC 720
QY 1026 GACCACTCGCTGTACACAGCAGCTCTGTGTATACACCCATCCGGGGAGTGGTATTAT 1085
Db 721 GACCACTCGCTGTACACAGCAGCTCTGTGTATACACCCATCCGGGGAGTGGTATTAT 780
QY 1086 GAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 1145
Db 781 GAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG 840
QY 1146 TACAACATATACAAGAGCATTTGGACAGTGGACACCAACCTTGGTTGGCCCAAGAAA 1205
Db 841 TACAACATATACAAGAGCATTTGGACAGTGGACACCAACCTTGGTTGGCCCAAGAAA 900
QY 1206 GTGTTTGAAGCTGCAGTCAAAATCATCAAGGCAGCCCTCTCCACGGAGAAAGTTCCCTGTAT 1265
Db 901 GTGTTTGAAGCTGCAGTCAAAATCCATCAAGCAGCCCTCTCCACGGAGAAAGTTCCCTGTAT 960
QY 1266 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGCACCCTTGGAAACATT 1325
Db 961 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGCACCCTTGGAAACATT 1020
QY 1326 TTCCCACTCATCTCACTCTACCTTAATGGGTGAGGTTACCAACCAAGCTCCTTCCGCATCAC 1385
Db 1021 TTCCCACTCATCTCACTCTACCTTAATGGGTGAGGTTACCAACCAAGCTCCTTCCGCATCAC 1080
QY 1386 ATCCTTCGGCAGCAATACCTTCGGCCAGTGGAAAGATGTGCGCCAGCTGCCAAGACGACTGT 1445
Db 1081 ATCCTTCGGCAGCAATACCTTCGGCCAGTGGAAAGATGTGCGCCAGCTGCCAAGACGACTGT 1140
QY 1446 TACAAGTTTGCCATCTCACAGTCAATCACCGGCGACTGTATGGGAGCTGTATCATCGAG 1505
Db 1141 TACAAGTTTGCCATCTCACAGTCAATCACCGGCGACTGTATGGGAGCTGTATCATCGAG 1200
QY 1506 GCGTCTACGTTGTCTTTGATCGGGCCGAAAAACGAATTCGCTTGTGTCAGCGCTTGC 1565
Db 1201 GCGTCTACGTTGTCTTTGATCGGGCCGAAAAACGAATTCGCTTGTGTCAGCGCTTGC 1260
QY 1566 CATGTGCACCATCAGTTCAGGAGCGGAGCGGTGGAGGCCCTTTTGTCACTTTGGACATG 1625
Db 1261 CATGTGCACCATCAGTTCAGGAGCGGAGCGGTGGAGGCCCTTTTGTCACTTTGGACATG 1320
QY 1626 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1685
Db 1321 GAAGACTGTGGCTTACAACTTCCACAGACAGATGAGTCAACCCCTCATGACCATAGCCCTAT 1380
QY 1686 GTCATGGCTGCCATCTGCGCCCTCTTCATCTGCGCACTCTGCCTCATGTGTGTGTCAGTGG 1745
Db 1381 GTCATGGCTGCCATCTGCGCCCTCTTCATCTGCGCACTCTGCCTCATGTGTGTGTCAGTGG 1440
QY 1746 CGCTGGCTCGCTGCTGCGCCAGCAGCATGATGACTTTTCTGATGACATCTCCCTGCTG 1805
Db 1441 CGCTGGCTCGCTGCTGCGCCAGCAGCATGATGACTTTTCTGATGACATCTCCCTGCTG 1500
QY 1806 AAGTGAAGGGCCCATGGCAGAAATAGAGATTCCTCCCT-GGACCAACACCTCCGTGGTTC 1864
Db 1501 AAGTGAAGGGCCCATGGGAGAAAGATAGAGATTCCTCCCTGGGACCAACCTCCGTGGTTC 1560
QY 1865 ACTTTGGTCAACAGTAGGAGACACAGATGCCACTGTGGCCAGAGCACTTCAGGACCCCTC 1924
Db 1561 ACTTTGGTCAACAGTAGGAGACACAGATGCCACTGTGGCCAGAGCACTTCAGGACCCCTC 1620
QY 1925 CCCACCCACCAATGCTCTGCTTGTATGGAGAGGAAAGGCTGGCAAGTGGTGGTTCCA 1984
Db 1621 CCCACCCACCAATGCTCTGCTTGTATGGAGAGGAAAGGCTGGCAAGTGGTGGTTCCA 1680
QY 1985 GGGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAGCACTCTGCTGGCGGGGAATAC 2044
Db 1681 GGGACTGTACCTGTAGGAAACAGAAAGAGAAAGAAAGCACTCTGCTGGCGGGGAATAC 1740
QY 2045 TCTTTGTCACTCAATTTTAAGTGGGAAATTTCTGCTGCTTGAACATTCAGCCCTGAACC 2104
|||||

Db 1741 TCTGTGTCACCTCAAAATTTAAGTCGGGAAATTTCTGTCTTGAACATTCAGCCCTGAACC 1800
QY 2105 TTTGTCCACCATTCCTTTAAATTTCTCAACCCCAAAAGTATTTCTTTTCTTCTTAGTTTCAGA 2164
Db 1801 TTTGTCCACCATTCCTTTAAATTTCTCAACCCCAAAAGTATTTCTTTTCTTCTTAGTTTCAGA 1860
QY 2165 AGTACTGGCATCACACGAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAA 2224
Db 1861 AGTACTGGCATCACACGAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAA 1920
QY 2225 GAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTAT 2284
Db 1921 GAGACCAAGCTTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGATGCACAGTTTGTCTAT 1980
QY 2285 TTGCTTTAGAGACAGGGACTGTATAACAAGCCCTAACATTTGGTCAAGATTTGCTCTTG 2344
Db 1981 TTGCTTTAGAGACAGGGACTGTATAACAAGCCCTAACATTTGGTCAAGATTTGCTCTTG 2040
QY 2345 AATT 2348
Db 2041 AATT 2044
RESULT 10
E50816
LOCUS Aspartate proteinase ASP2. 2541 bp DNA linear PAT 18-JUN-2001
DEFINITION Aspartate proteinase ASP2.
ACCESSION E50816
VERSION E50816.1 GI:13023199
KEYWORDS JP 2000060579-A/1.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2541)
AUTHORS David,J.P., Conrad,G.C., Kay,M. and Trudy,S.S.
TITLE Aspartate proteinase ASP2
JOURNAL Patent: JP 2000060579-A 1 29-FEB-2000;
SMITHKLINE BEECHAM CORP
COMMENT OS Unidentified
PN JP 2000060579-A/1
PD 29-FEB-2000
PF 03-AUG-1999 JP 1999219665
PI 28-JAN-1997 GB 9701684:4
PC C12N15/09,A61K31/7088,A61K38/46,A61K39/00,A61K39/395,PC
A61K39/395,A61K48/00,
PC A61P25/28,A61P35/00,A61P43/00,C07K16/40,C12N1/15,C12N1/19,PC
C12N1/21,
PC C12N5/10,C12N9/64,C12Q1/37,G01N33/15,G01N33/50,G01N33/53,PC
G01N33/566,
PC G01N33/577//C12P21/08,C12N15/00,A61K37/54,C12N5/00 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..2541
FT Location/Qualifiers
1..2541
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 598 a 673 c 675 g 579 t 16 others
ORIGIN
Query Match 86.2%; Score 2024; DB 6; Length 2541;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2038; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 306 ATGCCCAAGCCCTCGCCCTGCTGCTGTGGATGGCGGGAGTGTGCTTCCCTGCCAC 365
Db 1 ATGCCCAAGCCCTCGCCCTGCTGCTGTGGATGGCGGGAGTGTGCTTCCCTGCCAC 60
QY 366 GGCACCCAGCAGGCGATCCGCCCTCGCGAGCGCTGGGGGGCGCCCCCTGGG 425
|||||

REFERENCE	1 (bases 1 to 3252)
AUTHORS	Tang,J.-J., Hong,L. and Ghosh,A.K.
TITLE	Inhibitors of memapsin 2 and use thereof
JOURNAL	Patent: WO 0100665-A 1 04-JAN-2001; Oklahoma Medical Research Foundation (US) ; THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES	Location/Qualifiers
source	1..3252
BASE COUNT	804 a 863 c 811 g 771 t 3 others
ORIGIN	
Query Match	85.3%; Score 2004; DB 6; Length 3252;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2004; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	345 GCGGAGTGCCTGCCACGGACCCAGACAGGCGATCCGGCTGCCCTTCGCGAGCGC 404
Db	1 GCGGAGTGCCTGCCACGGACCCAGACAGGCGATCCGGCTGCCCTTCGCGAGCGC 60
Qy	405 CTGGGGGCGCCCCCTGGGGCTGCGGCTGCCCGGGAGACGACGAAGAGCCCGAGGAG 464
Db	61 CTGGGGGCGCCCCCTGGGGCTGCGGCTGCCCGGGAGACGACGAAGAGCCCGAGGAG 120
Qy	465 CCGGCCGGAGGGCAGCTTTTGTGGAGATGGTGGACACCTTGAGGGGCAAGTCGGGGCAG 524
Db	121 CCGGCCGGAGGGCAGCTTTTGTGGAGATGGTGGACAACTTGAGGGGCAAGTCGGGGCAG 180
Qy	525 GGCCTACTACGTGGAGATCACCGTGGGACGCCCGCGAGAGCGCTCAACATCCTGTGTGAT 584
Db	181 GGCCTACTACGTGGAGATCACCGTGGGACGCCCGCGAGAGCGCTCAACATCCTGTGTGAT 240
Qy	585 ACAGGCAGCAGTAACCTTTTGCACTGGGTGCTGCCCGCCACCCCTTCTCGATCGCTACTAC 644
Db	241 ACAGGCAGCAGTAACCTTTTGCACTGGGTGCTGCCCGCCACCCCTTCTCGATCGCTACTAC 300
Qy	645 CAGAGGCAGCTGTCCAGCACATACCGGACCTCCGGAGGGTGTGTATGTGCCCTTACACC 704
Db	301 CAGAGGCAGCTGTCCAGCACATACCGGACCTCCGGAGGGTGTGTATGTGCCCTTACACC 360
Qy	705 CAGGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCATGGCCCCCAAC 764
Db	361 CAGGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCATGGCCCCCAAC 420
Qy	765 GTCAGTGTGCGTGCCAAATGTCTGCCATCTACTGAATCAGACAAATTTCTTCAATCAACGC 824
Db	421 GTCAGTGTGCGTGCCAAATGTCTGCCATCTACTGAATCAGACAAATTTCTTCAATCAACGC 480
Qy	825 TCCAACTCGGAAGGCATCTCGGGCTGGCCTATGCTGAGATTGGCAGGCCTCAGCACATCC 884
Db	481 TCCAACTCGGAAGGCATCTCGGGCTGGCCTATGCTGAGATTGGCAGGCCTCAGCACATCC 540
Qy	885 CTGAGCCCTTTCTTTTGACTCTCTGGTAAGCAGACCCACGTTTCCCAACCTCTTCTCCCTG 944
Db	541 CTGAGCCCTTTCTTTTGACTCTCTGGTAAGCAGACCCACGTTTCCCAACCTCTTCTCCCTG 600
Qy	945 CAGCTTTGTGGTGTGGCTTCCCTCAACAGTCTGAAGTGTCTGGCTCTCTCGAGGG 1004
Db	601 CAGCTTTGTGGTGTGGCTTCCCTCAACAGTCTGAAGTGTCTGGCTCTCTCGAGGG 660
Qy	1005 AGCATGATCATGGAGGTATGCACACATCGCTGTACAGAGGAGTCTGTGGTATACACC 1064
Db	661 AGCATGATCATGGAGGTATGCACACATCGCTGTACAGAGGAGTCTGTGGTATACACC 720
Qy	1065 ATCCGGCGGGAGTGTATTATGAGTGTATCTTGTGCGGGTGGAGATCAATGGACAGGAT 1124
Db	721 ATCCGGCGGGAGTGTATTATGAGTGTATCTTGTGCGGGTGGAGATCAATGGACAGGAT 780
Qy	1125 CTGAAAATGGAGTGCAGAGGTACAACATATGACAGAGCAATTTGTGGACAGTGGCACACC 1184
Db	781 CTGAAAATGGAGTGCAGAGGTACAACATATGACAGAGCAATTTGTGGACAGTGGCACACC 840

Db 1621 GGTCGCAAGGTGGGTTCACGGAGCTGTACCTGTAGSAAACAGAAAAGAGAAAG 1680
QY 2025 CACTCTGCTGGCGGGAATACTCTTGTGTACCTCAAAATTTAAGTCGGGAAATTTCTGTGCT 2084
Db 1681 CACTCTGCTGGCGGGAATACTCTTGTGTACCTCAAAATTTAAGTCGGGAAATTTCTGTGCT 1740
QY 2085 TGAACCTTACAGCCCTGAACCTTTGTCCACCATTTCTTAAATTTCTCCAAACCCAAAGTATT 2144
Db 1741 TGAACCTTACAGCCCTGAACCTTTGTCCACCATTTCTTAAATTTCTCCAAACCCAAAGTATT 1800
QY 2145 CTTCCTTTCTTAGTTTTCAGAACTAGTGGCATACAGCAGGTTTACCTTGGCGTGTGTC 2204
Db 1801 CTTCCTTTCTTAGTTTTCAGAACTAGTGGCATACAGCAGGTTTACCTTGGCGTGTGTC 1860
QY 2205 TGTGTTACCTTGCAGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTTCAGTAGGA 2264
Db 1861 TGTGTTACCTTGCAGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTTCAGTAGGA 1920
QY 2265 GAGGATGCACAGTTTCTATTGCTTTTAGAGACAGGAGCTGTATAACAAAGCCCTAACATT 2324
Db 1921 GAGGATGCACAGTTTCTATTGCTTTTAGAGACAGGAGCTGTATAACAAAGCCCTAACATT 1980
QY 2325 GGTGCAAGATTGCCCTCTTGAATT 2348
Db 1981 GGTGCAAGATTGCCCTCTTGAATT 2004

RESULT 13
AX472368
LOCUS AX472368 3252 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 1 from Patent WO02053594.
ACCESSION AX472368
VERSION AX472368.1 GI:22207364
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Ghosh, A. K., Koelsch, G. and Tang, J. J.
TITLE Inhibitors of memapsin 2 and use thereof
PATENT: WO 02053594-A 1 11-JUL-2002;
JOURNAL OKLAHOMA MED RES FOUND (US); TRUSTEES OF THE UNIVERSITY OF ILLINOIS
(US)
FEATURES
Location/Qualifiers
1..3252
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 804 a 863 c 811 g 771 t 3 others
ORIGIN

Query Match 85.3%; Score 2004; DB 6; Length 3252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 GCGGGAGTGTGCTGCTGCCACGCCACCCACGAGCATCCGGCTGCCCTGCCAGCGGC 404
Db 1 GCGGGAGTGTGCTGCTGCCACGCCACCCACGAGCATCCGGCTGCCCTGCCAGCGGC 60
QY 405 CTGGGGGCCCCCTTGGGGCTGGGGCTGCCCGGGAGACCGACGAAGAGCCCGAGGAG 464
Db 61 CTGGGGGCGCCCTTGGGGCTGGGGCTGCCCGGGAGACCGACGAAGAGCCCGAGGAG 120
QY 465 CCGCGCGGAGGGGCGAGCTTTGTGGAGATGGTGACAACTTGGGGCAAGTCGGGGCAG 524
Db 121 CCGCGCGGAGGGGCGAGCTTTGTGGAGATGGTGACAACTTGGGGCAAGTCGGGGCAG 180
QY 525 GGCTACTACGTGGAGATGACCGTGGGGGAGCCCGCCGAGACGCTCAACATCTCTGGTGGAT 584
Db 181 GGCTACTACGTGGAGATGACCGTGGGGGAGCCCGCCGAGACGCTCAACATCTCTGGTGGAT 240
QY 585 ACAGGACAGTAACTTTTGCAGTGGGTGGTGGCCCCCACCCTTCCTGTCATCGCTACTAC 644
|||||

Db 241 ACAGGACGAGTAACATTTTGAGTGGGTGTGCCCCCACCCCTTCTCTGCATCGCTACTAC 300
QY 645 CAGAGCAGCTGTCCAGCACATACCGGAGCTCCGGAAGGSGTGTGTATGTGCCCCACACC 704
|||||
Db 301 CAGAGGAGCTGTCCAGCACATACCGGAGCTCCGGAAGGSGTGTGTATGTGCCCCACACC 360
QY 705 CAGGGCAAGTGGGAAGGGAGCTGGGCACCGACCTGTGTAAAGCATTCGCCCATGGCCCCAAC 764
|||||
Db 361 CAGGGCAAGTGGGAAGGGAGCTGGGCACCGACCTGTGTAAAGCATTCGCCCATGGCCCCAAC 420
QY 765 GTCACTGTGCTGCCCAACATTTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGC 824
|||||
Db 421 GTCACTGTGCTGCCCAACATTTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGC 480
QY 825 TCCAACTGGGAAGGCACTCTGGGCTGGCCATATGCTCAGATTGCGACGCTCTGACGACTCC 884
|||||
Db 481 TCCAACTGGGAAGGCACTCTGGGCTGGCCATATGCTCAGATTGCGACGCTCTGACGACTCC 540
QY 885 CTGAGGCTTTCTTTGACTCTCTGGTAAAGCAGACCCACCGTTCCCAACCTCTTCTTCCCTG 944
|||||
Db 541 CTGAGGCTTTCTTTGACTCTCTGGTAAAGCAGACCCACCGTTCCCAACCTCTTCTTCCCTG 600
QY 945 CAGCTTTGTGCTGCTGGCTTCCCTTCAACAGCTCTGAAGTGTGGCTCTGTGCGAGGG 1004
|||||
Db 601 CAGCTTTGTGCTGCTGGCTTCCCTTCAACAGCTCTGAAGTGTGGCTCTGTGCGAGGG 660
QY 1005 AGCATGATCAATTGGAGTATCGACCACTCGCTGTACACAGGCACTCTCTGTATACACC 1064
|||||
Db 661 AGCATGATCAATTGGAGTATCGACCACTCGCTGTACACAGGCACTCTCTGTATACACC 720
QY 1065 ATCCGGCGGAGTGGTATTATGAGGTGATCATTTGTGGGGTGGAGATCAATGGACAGGAT 1124
|||||
Db 721 ATCCGGCGGAGTGGTATTATGAGGTGATCATTTGTGCGGGTGGAGATCAATGGACAGGAT 780
QY 1125 CTGAAATGAGCTGCAAGAGGATCAACTATGACAAGAGCATTTGGGACAGTGGCACACC 1184
Db 781 CTGAAATGAGCTGCAAGAGGATCAACTATGACAAGAGCATTTGGGACAGTGGCACACC 840
QY 1185 AACCTTCTGTTGCCCAAGAAAGTGTTTGAAGTGCAGTCAAAATCCATCAAGGACGCTCC 1244
|||||
Db 841 AACCTTCTGTTGCCCAAGAAAGTGTTTGAAGTGCAGTCAAAATCCATCAAGGACGCTCC 900
QY 1245 TCCACGCGAGAAAGTTCCTGTATGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 1304
Db 901 TCCACGCGAGAAAGTTCCTGTATGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCA 960
QY 1305 GGCACACCCCTTGGAAACATTTTCCAGTCTATCTCACTCTACCTAATGGGTGAGGTTACC 1364
Db 961 GGCACACCCCTTGGAAACATTTTCCAGTCTATCTCACTCTACCTAATGGGTGAGGTTACC 1020
QY 1365 AACAGTCTTCCGCAATCAGCATCTTCCGAGCAATACCTTGGCGGCGAGTGGGAAGATGTG 1424
Db 1021 AACAGTCTTCCGCAATCAGCATCTTCCGAGCAATACCTTGGCGGCGAGTGGGAAGATGTG 1080
QY 1425 GCCACGTCCCAAGACGAGCTGTACAAAGTTTGCCATCTCACAGTCAATCCACGGGCACTGTT 1484
Db 1081 GCCACGTCCCAAGACGAGCTGTACAAAGTTTGCCATCTCACAGTCAATCCACGGGCACTGTT 1140
QY 1485 ATGGGAGCTGTTATCATGGAGGCTTCTACGTGCTCTTTGATGGGCCCCGAAACGAATT 1544
Db 1141 ATGGGAGCTGTTATCATGGAGGCTTCTACGTGCTCTTTGATGGGCCCCGAAACGAATT 1200
QY 1545 GGCCTTCTGTACGCGCTTGCATGTGCAGATGAGTTTACAGGCGGAGCGGTGGAAGGC 1604
Db 1201 GGCCTTCTGTACGCGCTTGCATGTGCAGATGAGTTTACAGGCGGAGCGGTGGAAGGC 1260
QY 1605 CCTTTTCTACCTTGGACATGGAGACTGGGTGTACAACTTCCACAGACAGATGAGTCA 1664
Db 1261 CCTTTTCTACCTTGGACATGGAGACTGGGTGTACAACTTCCACAGACAGATGAGTCA 1320
QY 1665 ACCCTCATGACCATAGGCTATGTATGGTGGCCATCTGTGGGCCCCCTTTCATGCTGCCACTC 1724
Db 1321 ACCCTCATGACCATAGGCTATGTATGGTGGCCATCTGTGGGCCCCCTTTCATGCTGCCACTC 1380
|||||

QY	1725	TGCTCATAGTGTGTCAAGTGGCGTGCCTCCGCTGCCTGCGCCAGCAGCATGACTTTT	1784
Db	1381	TGCTCATAGTGTGTCAAGTGGCGTGCCTCCGCTGCCTGCGCCAGCAGCATGACTTTT	1440
QY	1785	GCTGATGACATCTCCCTGCTCAAGTGGAGAGGCCCATGGCAGAGATGAGATTCCCTT	1844
Db	1441	GCTGATGACATCTCCCTGCTCAAGTGGAGAGGCCCATGGCAGAGATGAGATTCCCTT	1500
QY	1845	GGACCACACCTCCCTGGTTTCACTTTGGTTCACAAGTAGAGACACACATGCGACCTGTGGC	1904
Db	1501	GGACCACACCTCCCTGGTTTCACTTTGGTTCACAAGTAGAGACACACATGCGACCTGTGGC	1560
QY	1905	CAGAGCACCTCAGAGCCCTCCCCACCCACCAATGCCCTTGCTTGATGGAGAGAGAAAA	1964
Db	1561	CAGAGCACCTCAGAGCCCTCCCCACCCACCAATGCCCTTGCTTGATGGAGAGAGAAAA	1620
QY	1965	GGCTGGCAAGGTGGTTCAGGGACTGTACCTGTAGGAACAGAGAGAGAGAGAGAG	2024
Db	1621	GGCTGGCAAGGTGGTTCAGGGACTGTACCTGTAGGAACAGAGAGAGAGAGAGAG	1680
QY	2025	CACCTGCTCGCGGGAATACCTTTGGTTCACCTCAAAATTTAAGTCGGGAAATCTGCTGCT	2084
Db	1681	CACCTGCTCGCGGGAATACCTTTGGTTCACCTCAAAATTTAAGTCGGGAAATCTGCTGCT	1740
QY	2085	TGAACCTTACGCGCTGAACCTTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT	2144
Db	1741	TGAACCTTACGCGCTGAACCTTTTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATT	1800
QY	2145	CTTCTTTTCTTAGTTTTCAGAGTACTGGCATCACAGCAGGTTACCTTGGCGTGTGCC	2204
Db	1801	CTTCTTTTCTTAGTTTTCAGAGTACTGGCATCACAGCAGGTTACCTTGGCGTGTGCC	1860
QY	2205	TGTGGTACCTTGGCAGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA	2264
Db	1861	TGTGGTACCTTGGCAGAGAGAGACCAAGCTTGTTCCTGCTGGCCAAAGTCAGTAGGA	1920
QY	2265	GAGATGACACAGTTTGTATTGCTTTTAGAGACAGGAGTGTATTAACAAAGCCTAACATT	2324
Db	1921	GAGATGACACAGTTTGTATTGCTTTTAGAGACAGGAGTGTATTAACAAAGCCTAACATT	1980
QY	2325	GGTGCAGAGATTGCTCTTGAAAT 2348	
Db	1981	GGTGCAGAGATTGCTCTTGAAAT 2004	
RESULT 14			
AF200193			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1..3252			
/organism="Homo sapiens"			

/db_xref="taxon:9606"			
<1..1467			
/function="cleaves beta-secretase site of beta-amyloid precursor protein"			
/note="membrane-associated aspartic protease 2"			
/codon_start=1			
/product="memapsin 2"			
/protein_id="AAF13715.1"			
/db_xref="GI:6470293"			
/translation="AGVLPAGHTQHGIRLPLKSGLGAGPLGLRLPRETDEEPEEPPGR			
GSFVEMVDNLKCKSCOGYVYEMTVGSPPTLNIIVDTGSSNFACVAAAPPLHRYVOR			
QLSSYRDLRKGVYVPTQKWEGLGELDLVSIPIHPNVTVVRANTAAITESDKPFLING			
SNWEGILGLAYAEIARPDSDLEPFDSVLQKTHVPLNFSLQCLCGKAGFLPNQSEVLASV			
GSGMIIGDHLSTYGLSLWTYPIIRREMYEYIIVRVEINGDGLKMDCKEYNDKSIDV			
SCPTNLRLPKKVFEEAAVKSTKAASSTKFPDGLVQWQAGCTTPWNIIPVLSLY			
LMGEVTNQSFRTILPQYLRPVEDVATSDDCYKFAISQSSSTGTVMGAVIMEGFYVV			
FDRARKRIGFAVSACHVHDEFRTAAVEGPEVTLDMEDCGYNIPLQDTESTLMTIAYMA			
AICALFPLPLCLMVQWRCURLCRQQHDDFADDDISLLK"			
BASE COUNT 804 a 863 c 812 g 773 t			
ORIGIN			
Query Match 85.3%; Score 2004; DB 9; Length 3252;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	345	CGCGGAGTGTGCTGCCACCGCACCGACGCGCATCCGGCTGCCCTCGCGACGGC	404
Db	1	CGCGGAGTGTGCTGCCACCGCACCGACGCGCATCCGGCTGCCCTCGCGACGGC	60
Qy	405	CTGGGGGGGGCCCCCTGGGGCTGGCGGTGCCCGGGAGACGAGAGAGCCGAGGAG	464
Db	61	CTGGGGGGGGCCCCCTGGGGCTGGCGGTGCCCGGGAGACGAGAGAGCCGAGGAG	120
Qy	465	CCCGCGGGAGGGGCGAGCTTTGTGGAGATGTTGGACAACTTGAGGGGCAAGTCGGG	524
Db	121	CCCGCGGGAGGGGCGAGCTTTGTGGAGATGTTGGACAACTTGAGGGGCAAGTCGGG	180
Qy	525	GGCTACTAGTGGAGATGACCGTGGGACGCCCCCGCAGAGCGTCAACATCTCTGGT	584
Db	181	GGCTACTAGTGGAGATGACCGTGGGACGCCCCCGCAGAGCGTCAACATCTCTGGT	240
Qy	585	ACAGGCAGAGTAACCTTTGCAGTGGGTGCTGCCCGCCACCCCTTCTGCATCGCTAC	644
Db	241	ACAGGCAGAGTAACCTTTGCAGTGGGTGCTGCCCGCCACCCCTTCTGCATCGCTAC	300
Qy	645	CAGAGCAGCTGTCCAGACATACCCGGACCTCCGGGAAGGTGTGTATGTGCGCTACAC	704
Db	301	CAGAGCAGCTGTCCAGACATACCCGGACCTCCGGGAAGGTGTGTATGTGCGCTACAC	360
Qy	705	CAGGCAAGTGGGAAGGGAGCTGGGCACCGACCTTGGTAAGCATGCCCATATGCCCAAC	764
Db	361	CAGGCAAGTGGGAAGGGAGCTGGGCACCGACCTTGGTAAGCATGCCCATATGCCCAAC	420
Qy	765	GTCACTGTGGTGGCCCAACATTTGCTGCCATCACTGAATCAGACAACTTCTTCATCAAC	824
Db	421	GTCACTGTGGTGGCCCAACATTTGCTGCCATCACTGAATCAGACAACTTCTTCATCAAC	480
Qy	825	TCCAACTGGGAAGGCATCTCTGGGGCTGGCTTATGCTGAGATTGCCAGGCTTCACGACTCC	884
Db	481	TCCAACTGGGAAGGCATCTCTGGGGCTGGCTTATGCTGAGATTGCCAGGCTTCACGACTCC	540
Qy	885	CTGGAGCCCTTTCTTGTACTCTCTGGTAAAGCAGACCCACGTTGCCAACCTCTTCTCCCTG	944
Db	541	CTGGAGCCCTTTCTTGTACTCTCTGGTAAAGCAGACCCACGTTGCCAACCTCTTCTCCCTG	600
Qy	945	GAGCTTTGTGGTCTGGCTTCCCGCTCAACAGTCTGAAGTGTCTGGCTCTCTCGGAGGG	1004
Db	601	GAGCTTTGTGGTCTGGCTTCCCGCTCAACAGTCTGAAGTGTCTGGCTCTCTCGGAGGG	660
Qy	1005	AGCATCATTCATTCGAGGTATTCGACCACTCGCTGTACACAGGCGAGTCTCTGGTATACACC	1064
Db	661	AGCATCATTCATTCGAGGTATTCGACCACTCGCTGTACACAGGCGAGTCTCTGGTATACACC	720

```
QY 1065 ATCCGGGGGAGTGGTATATAGAGGTGATCAATTTGTCGGGTGGAGATCAATGACAGGAT 1124
Db 721 ATCCGGGGGAGTGGTATATAGAGGTGATCAATTTGTCGGGTGGAGATCAATGACAGGAT 780
QY 1125 CTCAAATGACATGCAAGGAGTACAACATATGACAAGACATTTGGACAGTGGCACACC 1184
Db 781 CTCAAATGACATGCAAGGAGTACAACATATGACAAGACATTTGGACAGTGGCACACC 840
QY 1185 AACCTTCGTTTGGCCCAAGAAAGTGTGTAAGCTGCAAGTCAAAATCCATCAAGGAGGCTCC 1244
Db 841 AACCTTCGTTTGGCCCAAGAAAGTGTGTAAGCTGCAAGTCAAAATCCATCAAGGAGGCTCC 900
QY 1245 TCACGGAGAAAGTTCCTTGATGTTTCTGGCTAGGAGACAGTGGTGTGCTGGCAAGCA 1304
Db 901 TCACGGAGAAAGTTCCTTGATGTTTCTGGCTAGGAGACAGTGGTGTGCTGGCAAGCA 960
QY 1305 GGCACACACCTTGGAAACATTTTCCAGTCATCTCACTTACCTAATGGGTGAGTTACC 1364
Db 961 GGCACACACCTTGGAAACATTTTCCAGTCATCTCACTTACCTAATGGGTGAGTTACC 1020
QY 1365 AACAGTTCCTTCCGCATCACCATCTTCCGCAGCAATACCTGCGGCCAGTGGGAAGATGTG 1424
Db 1021 AACAGTTCCTTCCGCATCACCATCTTCCGCAGCAATACCTGCGGCCAGTGGGAAGATGTG 1080
QY 1425 GCCACGTCCCAAGACGACTGTTTACAAGTTTGGCCATCTCACAGTCATCCACGGGCACCTGT 1484
Db 1081 GCCACGTCCCAAGACGACTGTTTACAAGTTTGGCCATCTCACAGTCATCCACGGGCACCTGT 1140
QY 1485 ATGGAGCTGTTATCATGGAGGCTTCTACGTTGTCTTTGATCGGGCCCCGAAACGAAT 1544
Db 1141 ATGGAGCTGTTATCATGGAGGCTTCTACGTTGTCTTTGATCGGGCCCCGAAACGAAT 1200
QY 1545 GGCCTTTCCTGTCAGCGCTGCCATGTGCAGATGAGTTTACAGGACGGCGGTGGGAAGGC 1604
Db 1201 GGCCTTTCCTGTCAGCGCTGCCATGTGCAGATGAGTTTACAGGACGGCGGTGGGAAGGC 1260
QY 1605 CTTTTCCTCACCCTGGACATGGAAGACTGTGGCTACAACATTCACACAGACAGATGAGTCA 1664
Db 1261 CTTTTCCTCACCCTGGACATGGAAGACTGTGGCTACAACATTCACACAGACAGATGAGTCA 1320
QY 1665 ACCCTCATGACCATAGCTATGTCATGGCTGCCATCTGCCCTCTTCATGCTGCCACTC 1724
Db 1321 ACCCTCATGACCATAGCTATGTCATGGCTGCCATCTGCCCTCTTCATGCTGCCACTC 1380
QY 1725 TGCTTCATGGTGTGTCAGTGGCGCTGCCTCCGCTGCTGCGCCAGCAGCATGATGACTTT 1784
Db 1381 TGCTTCATGGTGTGTCAGTGGCGCTGCCTCCGCTGCTGCGCCAGCAGCATGATGACTTT 1440
QY 1785 GCTGATGACATCTCCTGCTGAAGTGAAGAGGCCCATGGGCGAGAAGATAGAGATTCCTCT 1844
Db 1441 GCTGATGACATCTCCTGCTGAAGTGAAGAGGCCCATGGGCGAGAAGATAGAGATTCCTCT 1500
QY 1845 GGACACACCTCGTGGTTTACATTTGGTCAACAGTAGGACACACAGATGCACCTGTGGC 1904
Db 1501 GGACACACCTCGTGGTTTACATTTGGTCAACAGTAGGACACACAGATGCACCTGTGGC 1560
QY 1905 CAGAGCACCTTCAGGACCTCCCAACCCACCAATGCTCTGCTGCTGATGAGAAGGAAAA 1964
Db 1561 CAGAGCACCTTCAGGACCTCCCAACCCACCAATGCTCTGCTGCTGATGAGAAGGAAAA 1620
QY 1965 GGTGGCAAGTGGGTTCCAGGACTGTACTGTAGGAAACAGAAAAAGAGAAAGAAAG 2024
Db 1621 GGTGGCAAGTGGGTTCCAGGACTGTACTGTAGGAAACAGAAAAAGAGAAAGAAAG 1680
QY 2025 CACTTCTGCTGGGGGATACCTTGTGTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGT 2084
Db 1681 CACTTCTGCTGGGGGATACCTTGTGTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGT 1740
QY 2085 TGAATCTCAGCCCTTGAACCTTTGTCCACCATCTCTTAAATTTCTCCAAACCCAAAGTATT 2144
Db 1741 TGAATCTCAGCCCTTGAACCTTTGTCCACCATCTCTTAAATTTCTCCAAACCCAAAGTATT 1800
QY 2145 CTTTCTTTTCTTAGTTTTCAGAAAGTACTGGCATCACACGAGGTTTACCTTGGCGTGTGCC 2204
```

```
Db 1801 CTTCTTTCTTAGTTTTCAGAAAGTACTGGCATCACACGAGTTTACCTTGGCGTGTGCC 1860
QY 2205 TGTGGTACCCTGCAGAGAAGACCAAGCTTGTTCCTCTGCTGGCCAAAGTCAGTAGGA 2264
Db 1861 TGTGGTACCCTGCAGAGAAGACCAAGCTTGTTCCTCTGCTGGCCAAAGTCAGTAGGA 1920
QY 2265 GAGGATGCACAGTTTGTCTATTTCCTTTAGAGACAGGAGCTGTATAACAAGCCCTAACATT 2324
Db 1921 GAGGATGCACAGTTTGTCTATTTCCTTTAGAGACAGGAGCTGTATAACAAGCCCTAACATT 1980
QY 2325 GGTGCAAGATTGCCTCTTCAATT 2348
Db 1981 GGTGCAAGATTGCCTCTTCAATT 2004

RESULT 15
AX105387
LOCUS AX105387 1977 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 5 from Patent WO0123533.
ACCESSION AX105387
VERSION AX105387.1 GI:13921512
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1977)
AUTHORS Gurney,M. and Bienkowski,M.J.
TITLE Alzheimer's disease secretase, app substrates therefor, and uses
JOURNAL Patent: WO 0123533-A 5 05-APR-2001;
Pharmacia & Upjohn Company (US)
FEATURES
Location/Qualifiers
source 1..1977
/organism="Homo sapiens"
BASE COUNT 447 a 552 c 551 g 427 t
/db_xref="taxon:9606"
ORIGIN
Query Match 80.1%; Score 1881; DB 6; Length 1977;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1966; Conservative 0; Mismatches 0; Indels 75; Gaps 1;
QY 306 ATGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGCGGAGTGTGCTGCCCTGCCAC 365
Db 1 ATGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGCGCGGAGTGTGCTGCCCTGCCAC 60
QY 366 GGCACCCAGCACGGCATTCGGGCTGCCCTGCGCAGCGGCTTGGGGGCGCCCCCTGGGG 425
Db 61 GGCACCCAGCACGGCATTCGGGCTGCCCTGCGCAGCGGCTTGGGGGCGCCCCCTGGGG 120
QY 426 CTGGGCTGCCCCGGGAGACCGAGAGACCGCAGAGAGCCCGCAGAGAGCCCGCGGAGGGGAGCTTT 485
Db 121 CTGGGCTGCCCCGGGAGACCGAGAGAGCCCGCAGAGAGCCCGCGGAGGGGAGCTTT 180
QY 486 GTGGAGATGTGGCAACCTGAGGGGCAAGTTCGGGGCAGGCTACTACGTGGAGATGACC 545
Db 181 GTGGAGATGTGGCAACCTGAGGGGCAAGTTCGGGGCAGGCTACTACGTGGAGATGACC 240
QY 546 GTGGGACGCCCCCGCAGAGAGCTCAACATCTTGTGGATACAGGACAGCAGTAACTTTGCA 605
Db 241 GTGGGACGCCCCCGCAGAGAGCTCAACATCTTGTGGATACAGGACAGCAGTAACTTTGCA 300
QY 606 GTGGTCTGCCCCCGCAGAGAGCTTCCCTGCATCGCTTACCAGAGGACAGCTGTCCAGCACA 665
Db 301 GTGGTCTGCCCCCGCAGAGAGCTTCCCTGCATCGCTTACCAGAGGACAGCTGTCCAGCACA 360
QY 666 TACCGGACCTCGGAAGGTTGTGTATGTGCCCTTACACCCAGGCGCAAGTGGGAAGGGAG 725
Db 361 TACCGGACCTCGGAAGGTTGTGTATGTGCCCTTACACCCAGGCGCAAGTGGGAAGGGAG 420
QY 726 CTGGGCACCGACCTGGTGAAGCATCCCCCATGGCCCCCAACCGTCACTGTGCGTGTGCCAACATT 785
```

```
Db 421 CTGGGACCGACCTGGTAAAGCATCCCATGGCCCAAGCTCACTGTGCGTGCCAAACATT 480
QY 786 GCTCCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCCAACTGGGAAGGCATCCGT 845
Db 481 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCCAACTGGGAAGGCATCCGT 540
QY 846 GGGCTGGCCATATGCTAGATTGCCAGGCGCTGACGACTCCCTGGAGCCCTTTCTTTGACTCT 905
Db 541 GGGCTGGCCATATGCTAGATTGCCAG-----566
QY 906 CTGGTAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTC 965
Db 567 -----GCTTTGCTGGCTGGCTTC 585
QY 966 CCCCTCAACCACTCTGAAGTGTGGCTCTGTGCGGAGGAGCATGATCATTTGGAGGTATC 1025
Db 586 CCCCTCAACCACTCTGAAGTGTGGCTCTGTGCGGAGGAGCATGATCATTTGGAGGTATC 645
QY 1026 GACCACTCGCTGTACACAGGCACTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT 1085
Db 646 GACCACTCGCTGTACACAGGCACTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT 705
QY 1086 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGATCTGAATGGACTGCAAGGAG 1145
Db 706 GAGGTGATCATTTGCGGGTGGAGATCAATGGACAGATCTGAATGGACTGCAAGGAG 765
QY 1146 TACAACATATGACAAGAGCAATTTGGACAGTGGCCACCAACCACTTCTGTTGCCCAAGAAA 1205
Db 766 TACAACATATGACAAGAGCAATTTGGACAGTGGCCACCAACCACTTCTGTTGCCCAAGAAA 825
QY 1206 GTGTTTGAAGCTGCAGTCAATFCCATCAAGGAGCCTCTCTCCAGGAGAGTTCCTCGTAT 1265
Db 826 GTGTTTGAAGCTGCAGTCAATFCCATCAAGGAGCCTCTCTCCAGGAGAGTTCCTCGTAT 885
QY 1266 GGTTCCTGGCTAGGAGACAGCTGTGCTGCGCAAGCAGGACCAACCCCTTGGAAACATT 1325
Db 886 GGTTCCTGGCTAGGAGACAGCTGTGCTGCGCAAGCAGGACCAACCCCTTGGAAACATT 945
QY 1326 TTCCCACTCATCTCACTTACCTAATGGGTGAGGTTTACCAACCACTCTTCCCGCATCACC 1385
Db 946 TTCCCACTCATCTCACTTACCTAATGGGTGAGGTTTACCAACCACTCTTCCCGCATCACC 1005
QY 1386 ATCTTCCCGCAGCAATACCTCGGGCCAGTGGAGAGTGTGGCCACGTCGCCAAGACGACTGT 1445
Db 1006 ATCTTCCCGCAGCAATACCTCGGGCCAGTGGAGAGTGTGGCCACGTCGCCAAGACGACTGT 1065
QY 1446 TACAAGTTTGCATCTCACAGTCACTCCAGGCGCACTGTTATGGAGCTGTTATCATGGAG 1505
Db 1066 TACAAGTTTGCATCTCACAGTCACTCCAGGCGCACTGTTATGGAGCTGTTATCATGGAG 1125
QY 1506 GGCTTCTACGTTGCTTTGATCGGGCCCAACCAAGAAATGGCTTTGCTGTCAGGCGCTTGC 1565
Db 1126 GGCTTCTACGTTGCTTTGATCGGGCCCAACCAAGAAATGGCTTTGCTGTCAGGCGCTTGC 1185
QY 1566 CATGTGCAAGATAGTTTCAAGACGCGCAGGGGTGGAGGCGCTTTTGTCACTTGGACATG 1625
Db 1186 CATGTGCAAGATAGTTTCAAGACGCGCAGGGGTGGAGGCGCTTTTGTCACTTGGACATG 1245
QY 1626 GAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCAACCTCATGACCATAGACCTAT 1685
Db 1246 GAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCAACCTCATGACCATAGACCTAT 1305
QY 1686 GTCATGGCTGCGCATCTGGCCCTCTTTCATCTGCGCACTCTGCCTCATGTTGTGTGAGTGG 1745
Db 1306 GTCATGGCTGCGCATCTGGCCCTCTTTCATCTGCGCACTCTGCCTCATGTTGTGTGAGTGG 1365
QY 1746 CGCTGCGCTCGCTGCGTGGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1805
Db 1366 CGCTGCGCTCGCTGCGTGGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG 1425
QY 1806 AAGTGAGGAGGCCCATGGCAGAGATAGAGATTCCCTGGACCACTCCGTTGGTTCA 1865
|||||
```

```
Db 1426 AAGTGAGAGGCCCATGGGCGAAGATAGAGATTCCCTCTGGACACACACTCTCCGTGGTTTCA 1485
QY 1866 CTTTGGTCAACAAGTAGGAGACACAGATGCGCACCTGTGGCCAGAGCACTCAAGGACCCYCC 1925
Db 1486 CTTTGGTCAACAAGTAGGAGACACAGATGCGCACCTGTGGCCAGAGCACTCAAGGACCCYCC 1545
QY 1926 CCACCCACCAATGCTCTGCTTGTATGAGAGAGAAAGGCTGGCAAGGTGGGTTCCAG 1985
Db 1546 CCACCCACCAATGCTCTGCTTGTATGAGAGAGAAAGGCTGGCAAGGTGGGTTCCAG 1605
QY 1986 GGACTGTACCTGTAGGAAACAGAAAGAGAAAGAGCACTCTGCTGGCGGGAATACT 2045
Db 1606 GGACTGTACCTGTAGGAAACAGAAAGAGAAAGAGCACTCTGCTGGCGGGAATACT 1665
QY 2046 CTTGCTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 2105
Db 1666 CTTGCTCACCTCAAAATTTAAGTCGGGAAATTTCTGCTGCTTGAACCTTCAGCCCTGAACCT 1725
QY 2106 TTGTCCACCATTTCTTTAAATTTCTCCAAACCCAAAGATTTCTTTTCTTTAGTTTCAGAA 2165
Db 1726 TTGTCCACCATTTCTTTAAATTTCTCCAAACCCAAAGATTTCTTTTCTTTAGTTTCAGAA 1785
QY 2166 GTACTGGCATCACACGCGAGTTACCTTTGGCGTGTGTCCTGTCCTGTCACCTGGCAGAGAG 2225
Db 1786 GTACTGGCATCACACGCGAGTTACCTTTGGCGTGTGTCCTGTCCTGTCACCTGGCAGAGAG 1845
QY 2226 AGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCACTAGGAGAGGATGCACAGTTTGTCTATT 2285
Db 1846 AGACCAAGCTTTGTTTCCCTGCTGGCCAAAGTCACTAGGAGAGGATGCACAGTTTGTCTATT 1905
QY 2286 TGCTTTTAGAGACAGGGGACTGTATTAACAAGCCCTAACATTTGGTGCAGAAAGATTGCCTCTTGA 2345
Db 1906 TGCTTTTAGAGACAGGGGACTGTATTAACAAGCCCTAACATTTGGTGCAGAAAGATTGCCTCTTGA 1965
QY 2346 A 2346
Db 1966 A 1966
```

Search completed: March 2, 2003, 00:31:31
Job time : 4027 secs